

Db	1501	gggattgtccaaggggtcccccggcgcccacagtcagtgagtgagcagagatgtccac	1506
QY	1756	tatgagtcacaacaacagatctacacatgagtgagtcgctcactctctatagtcgcccacg	1815
Db	1561	tatgacgtacaacagcagatctacaacatgagtgagtcgcccgcactctccatagtgctgcacatg	1620
QY	1816	tccatcccgacaacacacacacatgtgcactctgtagagttagtgctccacagcccgacaacacatg	1875
Db	1621	tccacatcccgacaacacacacacatgtgcactctgtagagttagtgctccacagcccgacaacacatg	1668
QY	1876	cgacatgacagacacacagagaaaccaaataagctcacgctc---ccgagcacacagtgagctg	1932
Db	1669	cggactcgtacagacacacagagaaaccaaataagttgtccctgcgcccgcgcccgcgcgagtgct	1728
QY	1933	ttcccccctgtctctgagagaagttgtgtgaacacccggtcacagc-ctccctgtctatggaa	1991
Db	1729	ttcccccctgtctctgagagaagttgtgtgaacacccggtcacagcctctctgtctatggaa	1786
QY	1992	cttgagcaaatcttggtatcacagatgtgcgcgtgcctccacagtaacccacacacaccttg	2051
Db	1787	cggagacaacatctttctatcacagatgtgcgcgcgtgcctccaca---ccctaaccccacatcccg	1843
QY	2052	gtcccgagcgcgctgtgtgcctgcacatactatcatatgagatccctgtgtgggacgcccata-	2109
Db	1844	gcccctgagcgtgtgtgcgcgcgcacatatctatcaaaaatacatgtgtgtgggagccctgtcc	1903
QY	2110	ccgtgtctgtactctccactctgacctgagcttgacttgacatctgctggaagagagcaatctcc	2169
Db	1904	cccccctctgcgcgcctctacacacgtacacctgagctgtgtcatctgtctggaacagcgccatg	1963
QY	2170	cccccaacccccacagctgtgacacctgacacagagcagagagagcgacagctgtccacacaaga	2229
Db	1964	gctctgcacagcctgtgcctgc------	1982
QY	2230	cctgtgcagacacacacacacacacccgtctctgtgcacacccctccactacacctgagtgacagac	2289
Db	1983	-----caggtccctctagacacctgtcccccctgctgtctccagtgagagagttagcctg	2034
QY	2290	agccagctgagacctctcctctcaactacacagcgacacatggtcacatgagcgtgtacatgt	2349
Db	2035	gcccagcgagcggtccctccctctgcagcagacagcgctc---ggtcacacaagacgtatgacatgt	2092
QY	2350	gcttttttaattatattttttttagaagaaagacacggtgtcaacccacacagacctctgag	2409
Db	2093	gcttttttaattatattttttttagaagaaagacacggtgtcaactccgcagacacctctgtg	2152
QY	2410	aaacccgctgtg-cgcgcacaaagcagacacacccctgtctctctagggccacagag-cttcatg	2467
Db	2153	aagccagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcctctcctctatgactagagtgcgcgcg	2212
QY	2468	gaggggtgagccctgtcaagcctctcacagatgtgggcacacgcccctccacccaagaggtctaac	2527
Db	2213	gaggggtgagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcctccacccaagaggtctaac	2266
QY	2528	tcaaacctggaatgtacaacacacccacacgcctgtccaagcgc 2567	
Db	2267	tcaacatctggaatgtacaacacacccacacgcctgtccaagcgc 2306	

RESULT	2
246024	
ID	246024 standard; cDNA; 2223 BP.
XX	
AC	246024;
XX	
DT	25-APR-2000 (first entry)
XX	
DE	cDNA encoding a human transferase designated H07RAN-3.
XX	
KW	Transferase; H07RAN-3; arginine methyltransferase; autoimmune disorder;
KW	Inflammatory disorder; AIDS; atherosclerosis;
KW	adult respiratory distress syndrome; allergy; asthma; trauma;

KW autoimmune thyroiditis; bronchitis; Crohn's disease; diabetes mellitus;
KW gout; Grave's disease; osteoarthritis; osteoporosis; pancreatitis;
KW psoriasis; rheumatoid arthritis; infection; neurological disorder;
KW epilepsy; ischemic cerebrovascular disease; stroke; Alzheimer's disease
KW Picks disease; Huntington's disease; dementia; Parkinson's disease;
KW extrapyramidal disorder; Viral central nervous system disease;
KW prion disease; central nervous system developmental disorder;
KW neuroskeletal disorder; muscular dystrophy; neuromuscular disorder;
KW peripheral nervous system disorder; mental disorder; schizophrenic;
KW anxiety; reproductive disorder; gastrointestinal disorder; dyspepsia;
KW indigestion; gastritis; anorexia; nausea; abdominal angina;
KW gastroenteritis; intestinal obstruction; peptic ulcer;
KW irritable bowel syndrome; diarrhoea; constipation;
KW gastrointestinal haemorrhage; cancer; ss.

XX	OS	XX	FM	Key	Location/Qualifiers
----	----	----	----	-----	---------------------

FT	CDS
FT	224..1566
FT	/*tag= a
FT	/product= "transferase"
FT	/transl_except= {pos: 1100..1101, aa: Ala}

sig_peptide

XX
PN WO200000594-A2.

XX 06-JAN-2000.
PD

XX	29-JUN-1999;	99WO-US14651.
PF		

XX 30-JUN-1998; 98US-0109204.
PR

XX
PA (INCY-) INCYTE PHARM INC.

XX
PI Lal P, Bandman O, Hillman JL, Guegler KJ, Gorgone GA, Corley NC.

PI Patterson C;
XX

DR WPI; 2000-147267/13.
DR P-PSDB; Y54593.

XX Novel human transferases used for the diagnosis, treatment, and
PT

PT prevention of autoimmune/inflammatory, neurological, reproductive and
PT gastrointestinal disorders and cancer -

XX	Claim 7; Page 82; 95pp; English
PS	

XX The present sequence encodes a transferase polypeptide, designated

CC HUTRAN-3. The HUTRAN cDNA sequence is derived from Incyte clone
CC 2525071. The sequence is obtained from the partial sequences 246039-47
CC

CC The polypeptide is a arginine methyltransferase. The HMTAN polypeptides,
CC polynucleotides, agonists, antagonists, and antibodies can be used to

CC diagnose, treat or prevent autoimmune/inflammatory diseases (e.g. AIDS,
CC adult respiratory distress syndrome, allergies, asthma, atherosclerosis,
CC

CC autoimmune thyroiditis, Crohn's disease, diabetes mellitus, gout, Grave's disease, osteoarthritis, osteoporosis, pancreatitis, CC

CC psoriasis, rheumatoid arthritis, infections, trauma, neurological disorders (e.g. epilepsy, ischemic cerebrovascular disease, stroke, CC

CC Alzheimer's disease, Pick's disease, Huntington's disease, dementia,
CC Parkinson's disease and other extrapyramidal disorders, viral central
CC

CC nervous system disease, prion diseases, central nervous system
CC developmental disorders, neuroskeletal disorders, muscular dystrophy,
CC

neuromuscular disorders, peripheral nervous system disorders, mental disorders, schizophrenic, anxiety, reproductive disorders and

CC gastrointestinal disorders (e.g. dyspepsia, indigestion, gastritis, anorexia, nausea, abdominal angina, gastroenteritis, intestinal

CC obstruction, intestinal tract infections, peptic ulcer, irritable bowel
CC syndrome, diarrhoea, constipation, gastrointestinal haemorrhage, and
CC

CC cancer, XX

50 Sequence 2223 BP; 465 A; 732 C; 580 G; 444 T; 2 other;

Query Match 47.3%; Score 1478; DB 21; Length 2223;
Best Local Similarity 82.6%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 311; Indels 88; Gaps 14;

QY 290 cagaatgcatgtgtgtgacagacagatctctcaatcaaccctgggtcgaacagctgc 349
Db 1 cagaatgcatgtgtgtgacagacagatctctcaatcaaccctgggtcgaacagctgc 60
QY 350 tcaatcaatgt 407
Db 61 tcaatcaatgt 120
QY 408 tgggggccaac 467
Db 121 cgggggccaac 180
QY 468 gcaagcatctccagctctcaatcaatcaatcaatcaatcaatcaatcaatcaatca 527
Db 181 gcaagcatctccagctctcaatcaatcaatcaatcaatcaatcaatcaatcaatca 240
QY 528 gcgagacagac 587
Db 241 gcgagacagac 300
QY 588 cgttctagatgt 647
Db 301 cgttctagatgt 360
QY 648 caggaataatltatgt 707
Db 361 caggaataatltatgt 420
QY 708 gactaacaatctgt 767
Db 421 gactaacaatctgt 480
QY 768 gcttgaagacagtggaatatactcaatcaatcaatcaatcaatcaatcaatcaat 827
Db 481 gcttgaagacagtggaatatactcaatcaatcaatcaatcaatcaatcaatcaat 540
QY 828 aatgtctgt 887
Db 541 aatgtctgt 600
QY 888 caacatgt 947
Db 601 caacatgt 660
QY 948 caacaaagac 1007
Db 661 caacaaagac 720
QY 1008 aggt 1067
Db 721 aggt 780
QY 1068 cctgt 1127
Db 781 cctgt 840
QY 1128 gcaacagatagaatccatcaatcaatcaatcaatcaatcaatcaatcaatcaatca 1187
Db 841 gcaacagatagaatccatcaatcaatcaatcaatcaatcaatcaatcaatcaatca 900
QY 1188 ggccttctgt 1247
Db 901 ggccttctgt 960
QY 1248 ccccaacagagccctgt 1307
Db 961 cccgaacagagccctgt 1020
QY 1308 tgcgaagcgccgggacacgctctcagtgatgtctgtgtgtgtgtgtgtgtgtgtgt 1367

Db 1021 cgcgaagcgccgggacacgctctcaccaggaatgtgtgtgtgtgtgtgtgtgtgtgtgt 1080
QY 1368 ctatgacatcagatgt 1427
Db 1081 ctatgacatcagatgt 1139
QY 1428 ggtatcaagaac 1487
Db 1140 ggtatcaagaac 1199
QY 1488 acacatcagctctccctcggagatatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1547
Db 1200 acacatcagctctccctcggagatatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1259
QY 1548 cgggt 1607
Db 1260 cgggt 1319
QY 1608 ctccagtggt 1667
Db 1320 ctccagtggt 1379
QY 1668 ccaatcccgatgt 1727
Db 1380 ccaatcccgatgt 1439
QY 1728 ggt 1784
Db 1440 ggt 1499
QY 1785 tggccctgt 1844
Db 1500 tggccctgt 1559
QY 1845 gactgaagtgctgt 1904
Db 1560 gactgaagtgctgt 1607
QY 1905 agtccagagc---ccgacacagacagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1961
Db 1608 agtccagagc---ccgacacagacagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1665
QY 1962 caccgggt 2020
Db 1666 caccgggt 1725
QY 2021 ctgcccctcaagtaacccac 2080
Db 1726 ctgcccctcaagtaacccac 1782
QY 2081 acatgtatcctgt 2138
Db 1783 acatgtatcctgt 1842
QY 2139 cttgtatcctgt 2198
Db 1843 cttgtatcctgt 1892
QY 2199 gcaagagagcgacagctgt 2258
Db 1893 gcaagagagcgacagctgt 1913
QY 2259 ttgtacatccctcaatcctgt 2318
Db 1914 ttgtacatccctcaatcctgt 1973
QY 2319 ggcacatgt 2378
Db 1974 ggcacatgt 2031
QY 2379 agaacagtggtcaac 2437

Db 2032 agaacagtgatcgaacacgacccctctgtgaagccgagccgagccgagccagcag 2091
 Qy 2438 cccctgttcttagcagccagag-ttctaggtgaggggtgccccttcaagcttcagagt 2496
 Db 2092 cccctctccttagctcagagcgagcgaggaggggtgcccgcgaggttcagg- 2150
 Qy 2497 gggcagagccctccaccagaggttaccctcaacttgatgtacaaaccagcact 2556
 Db 2151 -----gccctcccnaccaactgttaccctcacttgatgtacaaaccagcact 2205
 Qy 2557 gtccaagggc 2567
 Db 2206 gtccggagagc 2216
 RESULT 3
 ID 246042 standard; cDNA; 621 BP.
 AC 246042;
 XX 25-APR-2000 (first entry)
 DE Partial cDNA encoding a transferase designated HUTRAN-3 fragment.
 KW Transferase; HUTRAN-3; arginine methyltransferase; autoimmune disorder;
 KW inflammatory disorder; AIDS; atherosclerosis;
 KW adult respiratory distress syndrome; allergy; asthma; trauma;
 KW autoimmune thyroiditis; bronchitis; Crohn's disease; diabetes mellitus;
 KW gout; Grave's disease; osteoarthritis; osteoporosis; pancreatitis;
 KW psoriasis; rheumatoid arthritis; infection; neurological disorder;
 KW epilepsy; ischemic cerebrovascular disease; stroke; Alzheimer's disease;
 KW picks disease; Huntington's disease; dementia; Parkinson's disease;
 KW extrapyramidal disorder; viral central nervous system disease;
 KW prion disease; central nervous system developmental disorder;
 KW neuroskeletal disorder; muscular dystrophy; neuromuscular disorder;
 KW peripheral nervous system disorder; mental disorder; schizophrenia;
 KW anxiety; reproductive disorder; gastrointestinal disorder; dyspepsia;
 KW indigestion; gastritis; anorexia; nausea; abdominal angina;
 KW gastroenteritis; intestinal obstruction; peptic ulcer;
 KW irritable bowel syndrome; diarrhoea; constipation;
 KW gastrointestinal haemorrhage; cancer; ss.
 XX Homo sapiens.
 OS
 XX MO200000594-A2.
 XX 06-JAN-2000.
 XX PD
 XX 29-JUN-1999; 99MO-US14651.
 XX PF
 XX 30-JUN-1998; 98US-0109204.
 XX PR
 XX (INCYTE PHARM INC.
 PA
 XX Lal P, Bandman O, Hillman JL, Guegler KJ, Gorgone GA, Corley NC;
 PI Patterson C;
 XX
 DR WPI; 2000-147267/13.
 XX
 PT Novel human transferases used for the diagnosis, treatment, and
 prevention of autoimmune/inflammatory, neurological, reproductive and
 gastrointestinal disorders and cancer
 XX
 PS Disclosure; Page 90; 95pp; English.
 XX
 CC 246039-47 represent partial fragments which were used to produce Incyte
 CC clone 2525071, which encodes a transferase polypeptide, designated
 CC HUTRAN-3. The polypeptide is an arginine methyltransferase. The
 CC HUTRAN polypeptides, polynucleotides, agonists, antagonists, and
 CC antibodies can be used to diagnose, treat or prevent autoimmune/
 CC inflammatory diseases (e.g. AIDS, adult respiratory distress syndrome,
 CC allergies, asthma, atherosclerosis, autoimmune thyroiditis, bronchitis,

CC Crohn's disease, diabetes mellitus, gout, Grave's disease,
 CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, rheumatoid
 CC arthritis, infections, trauma, neurological disorders (e.g. epilepsy,
 CC ischemic cerebrovascular disease, stroke, Alzheimer's disease, picks
 CC disease, Huntington's disease, dementia, Parkinson's disease and other
 CC extrapyramidal disorders, viral central nervous system disease, prion
 CC diseases, central nervous system developmental disorders, neuroskeletal
 CC disorders, muscular dystrophy, neuromuscular disorders, peripheral
 CC nervous system disorders, mental disorders, schizophrenia, anxiety,
 CC reproductive disorders and gastrointestinal disorders (e.g. dyspepsia,
 CC indigestion, gastritis, anorexia, nausea, abdominal angina,
 CC gastroenteritis, intestinal obstruction, intestinal tract infections,
 CC peptic ulcer, irritable bowel syndrome, diarrhoea, constipation,
 CC gastrointestinal haemorrhage, and cancer.
 XX
 SO Sequence 621 BP; 135 A; 180 C; 150 G; 140 T; 16 other;
 Query Match 15.8%; Score 493; DB 21; Length 621;
 Best Local Similarity 88.0%; Pred No. 1.7e-115;
 Matches 529; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 Qy 706 aagagtaacaatctgacagacgcatcgtgtatcccttgcgcaagtagagagttca 765
 Db 21 atgcmaattagagcncgagcccaacgctcctccgaggtgagaggtgca 80
 Qy 766 ttgcttgagcaagtggacattatcatctagaagccatgggtacatgcttcaatgaa 825
 Db 81 cttcccgagcaggttgcacatcatctcgtgagccatgggtacatgcttcaacagag 140
 Qy 826 cgaatgctgagagctacacctcagcagcaaaagttacttaagctgttgaacacgttc 885
 Db 141 cgcagctgganagctactctcagcagcaagagttacttgaagccagcggaacatgctt 200
 Qy 886 ccacacattgtatgtccacctgcaccccttcaactatgatacagctcacaatgagcag 945
 Db 201 cctacattgtgacgtccacctgcaccccttcaactatgatacagctcacaatgagcag 260
 Qy 946 ttccacaaagcacttcgagcagcagcagcagcagcagcagcagcagcagcagcagc 1005
 Db 261 ttccacaaagcacttcgagcagcagcagcagcagcagcagcagcagcagcagcagc 320
 Qy 1006 aagagtgccgctgtgagtagtacttcgagcagcagcagcagcagcagcagcagcag 1065
 Db 321 cgaggtgcgcgctgnt 380
 Qy 1066 atcctgatggcacaatctgtcaagtagacagtgaaacttcttgaagccaaagagcgat 1125
 Db 381 atcctgatggcacaatctgtcaagtagacagtgaaacttcttgaagccaaagagat 440
 Qy 1126 ttgcacagagataaataccatcaatcaatcaatcaatcaatcaatcaatcaatcaat 1185
 Db 441 ttgcacagagataaataccatcaatcaatcaatcaatcaatcaatcaatcaatcaat 500
 Qy 1186 ttggcctctgttgcagtggttcttcatgttgcctcacaataacgttgggtatccaca 1245
 Db 501 ctgagcttcctgttgcagtggttcttcatgttgcctcacaataacgttgggtatccaca 560
 Qy 1246 gccccaagagccctgacacacatgtagcagcagcagcagcagcagcagcagcagcag 1305
 Db 561 gccccaagagccctgacacacatgtagcagcagcagcagcagcagcagcagcagcag 620
 Qy 1306 t 1306
 Db 621 t 621
 RESULT 4
 ID 246043 standard; cDNA; 549 BP.
 AC 246043;
 XX
 XX

25-APR-2000 (first entry)

Partial cDNA encoding a transferase designated HUTRAN-3 fragment.

Transferase; HUTRAN-3; arginine methyltransferase; autoimmune disorder; inflammatory disorder; AIDS; atherosclerosis; adult respiratory distress syndrome; allergy; asthma; trauma; autoimmune thyroiditis; bronchitis; Crohn's disease; diabetes mellitus; gout; Grave's disease; osteoarthritis; osteoporosis; pancreatitis; psoriasis; rheumatoid arthritis; infection; neurological disorder; epilepsy; ischemic cerebrovascular disease; stroke; Alzheimer's disease; Picks disease; Huntington's disease; dementia; Parkinson's disease; extrapyramidal disorder; viral central nervous system disease; prion disease; central nervous system developmental disorder; neuroskeletal disorder; muscular dystrophy; neuromuscular disorder; peripheral nervous system disorder; mental disorder; schizophrenia; anxiety; reproductive disorder; gastrointestinal disorder; dyspepsia; indigestion; gastritis; anorexia; nausea; abdominal angina; gastroenteritis; intestinal obstruction; intestinal tract infections; peptic ulcer; irritable bowel syndrome; diarrhoea; constipation; gastrointestinal haemorrhage; and cancer.

Sequence 549 BP; 97 A; 144 C; 182 G; 126 T; 0 other;

Query Match 14.28; Score 443; DB 21; Length 549;
 Best Local Similarity 89.48; Pred. No. 7.9e-103;
 Matches 489; Conservative 0; Mismatches 55; Indels 3; Gaps 1;

1230 cgtgtgctatccacagcccaacagagcccttaccactggtacaggtccgtgct 1289
 |||||
 549 CGTGTGCTGTCACAGCCCGACAGAGCCCTTGACCCACATGTAACGATGCTGCT 490

1290 ctccagtcacacctgtgtttccagagccgggagacagctcccaaggatctcttat 1349
 |||||
 489 GTTCAGTCACCACTGTTGCCAAGCGGACAGCTCTCAGGAGATGCTCTTAT 430

1350 tgcacaacaaagacagagctatgacatcagttgtgacaggtggaacagagctc 1409
 |||||
 429 TGCACAACAAAGACAGAGCTACGACATCATGTTGTGCCAGGTGACACAGCCGCTC 370

1410 caagtcagtracactgtcgtgattctaaagacccctcttcaggtacagagtracaccc 1469
 |||||
 369 CAAGTCCAGTAACCTCTGATCTGAATAAACCCCTCTTTAGATACAGGGACACAGCC 310

1470 atcaccccccactgtgtctacactacagctctccctcggaagataatgtgaacacagaag 1529
 |||||
 309 CTCACCCCCACCCGGCTCCACTACACATCTCCCTCGGAAAAACATGTGAACACGGCGAG 250

1530 caactataatccagcagcgggtgtgctgtgtggaatgctactgctcagactgag 1589
 |||||
 249 CACCTACAACTCAGACCGGGAGTGGCCGACAGGATGCCGACCGCTATGACTTGA 190

1590 cagtgtatgtccggtccgctccaggtgtgtgtacacacacactgattcccttagtaaac 1649
 |||||
 189 CAGTGTATTTCAGTAGGCTCCAGCGTGGCCACACAACTGATTCCTTTGCCACACAC 130

1650 agggattgtcaatcacacacccactccggatgtgtgtcctaataatgagcagggatgtcca 1709
 |||||
 129 GGGGATTGTCAATCACACCACTCCCGGATGGGCTCCAAATATGACACAGGGGATGTCCA 70

1710 aggtctctcaggtgtcccaag---gaggcgggtgtgtgtccagttgcccactatgca 1766
 |||||
 69 AGGGTCTCTCCGGCCGACGAGGACATGTGTGGACAGAGTATGCCCATATGACATGCA 10

1767 caacag 1773
 |||||
 9 CAGCCAG 3

RESULT 5
 ID Z46044 standard; cDNA; 647 BP.
 AC Z46044;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Partial cDNA encoding a transferase designated HUTRAN-3 fragment.
 XX
 XX Transferase; HUTRAN-3; arginine methyltransferase; autoimmune disorder; inflammatory disorder; AIDS; atherosclerosis; adult respiratory distress syndrome; allergy; asthma; trauma; autoimmune thyroiditis; bronchitis; Crohn's disease; diabetes mellitus; gout; Grave's disease; osteoarthritis; osteoporosis; pancreatitis; psoriasis; rheumatoid arthritis; infection; neurological disorder; epilepsy; ischemic cerebrovascular disease; stroke; Alzheimer's disease; Picks disease; Huntington's disease; dementia; Parkinson's disease; extrapyramidal disorder; viral central nervous system disease; prion disease; central nervous system developmental disorder; neuroskeletal disorder; muscular dystrophy; neuromuscular disorder; peripheral nervous system disorder; mental disorder; schizophrenia; anxiety; reproductive disorder; gastrointestinal disorder; dyspepsia; indigestion; gastritis; anorexia; nausea; abdominal angina; gastroenteritis; intestinal obstruction; peptic ulcer; irritable bowel syndrome; diarrhoea; constipation; gastrointestinal haemorrhage; and cancer.

Homo sapiens.
 OS
 XX
 PN WO200000594-A2.

XX 06-JAN-2000.
 XX 29-JUN-1999; 99MO-US14651.
 XX 30-JUN-1998; 98US-0109204.
 XX (INCY-) INCYTE PHARM INC.
 XX Lal P, Bandman O, Hillman JL, Guegler KJ, Gorgone GA, Corley NC,
 PI Patterson C;
 XX WPI: 2000-147267/13.
 XX Novel human transferases used for the diagnosis, treatment, and
 PT prevention of autoimmune/inflammatory, neurological, reproductive and
 PT gastrointestinal disorders and cancer
 XX Disclosure: Page 91; 95pp; English.

XX Z46039-47 represent partial fragments which were used to produce Incyte
 CC clone 2525071, which encodes a transferase polypeptide, designated
 CC HUTRAN-3. The polypeptide is an arginine methyltransferase. The
 CC HUTRAN polypeptides, polynucleotides, agonists, antagonists, and
 CC antibodies can be used to diagnose, treat or prevent autoimmune/
 CC inflammatory diseases (e.g. AIDS, adult respiratory distress syndrome,
 CC allergies, asthma, atherosclerosis, autoimmune thyroiditis, bronchitis,
 CC Crohn's disease, diabetes mellitus, gout, Grave's disease,
 CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, rheumatoid
 CC arthritis, infections, trauma, neurological disorders (e.g. epilepsy,
 CC ischemic cerebrovascular disease, stroke, Alzheimer's disease, Pick's
 CC disease, Huntington's disease, dementia, Parkinson's disease and other
 CC extrapyramidal disorders), viral central nervous system disease, prion
 CC diseases, central nervous system developmental disorders, neuroskeletal
 CC disorders, muscular dystrophy, neuromuscular disorders, peripheral
 CC nervous system disorders, mental disorders, schizophrenia, anxiety,
 CC reproductive disorders and gastrointestinal disorders (e.g. dyspepsia,
 CC indigestion, gastritis, anorexia, nausea, abdominal angina,
 CC gastroenteritis, intestinal obstruction, intestinal tract infections,
 CC peptic ulcer, irritable bowel syndrome, diarrhoea, constipation,
 CC gastrointestinal haemorrhage, and cancer.

XX Sequence 647 BP; 142 A; 223 C; 167 G; 113 T; 2 other;

Query Match 13.4%; Score 420; DB 21; Length 647;
 Best Local Similarity 83.5%; Pred. No. 5.9e-97;
 Matches 541; Conservative 0; Mismatches 87; Indels 20; Gaps 5;

QY 1392 ggtggaccagagagctcgaagtcgaactgtgctgataaagaacccctcttcag 1451
 DB 1 ggtggaccagagagctcgaagtcgaactgtgctgataaagaacccctcttcag 60
 QY 1452 gtacacaggtacacacccatacccccacctgtgtccacatacagctctccctcgagaa 1511
 DB 61 atacacgggcacacagcctccaccccccacccgctccacatacactctccctcgagaa 120
 QY 1512 tatgtggacacaggaagacataatctagcagcggggtgtgtgtgtgtgtgtgtgtgtgc 1571
 DB 121 catgtggacacagggagacacatacactcagcagcgagcggtgtgtgtgtgtgtgtgtgtgc 180
 QY 1572 tactgtcagacagcagctgtattatgtccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgc 1631
 DB 181 gacgcctatgactgtgagcaggtgtattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgc 240
 QY 1632 gattcccttaacacagggattgtcaatacaacacacacccctccgattgtgtgtgtgtgtgc 1691
 DB 241 gattcccttaacacagggattgtcaatacaacacacacccctccgattgtgtgtgtgtgtgc 300
 QY 1692 gagcagcgagctgtgtcagagctcctcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgc 1748
 DB 301 gagcagcgagctgtgtcagagctcctcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgc 360

QY 1749 tgcacatagcagtcacacacacagtcacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgc 1807
 DB 361 tgcacatagcagtcacacacacagtcacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgc 420
 QY 1808 cgcacatgtcattccgcac 1867
 DB 421 cgcacatgtcattccgcac 469
 QY 1868 cagcagctgcagctgacagac 1924
 DB 470 cgcacagcgagac 529
 QY 1925 -agtgc 1982
 DB 530 gggcaggtgc 589
 QY 1983 ctatgggaactgtgacaatttgtacacagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgc 2030
 DB 590 ctatgggaactgtgacaatttgtacacagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgc 637

RESULT 6

V88444 ID V88444 standard; cDNA; 504 BP.

V88444; AC V88444;

DT 12-FEB-1999 (first entry)

XX EST clone G128.

XX Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

XX Homo sapiens.

XX W09845437-A2.

PD 15-OCT-1998.

PF 10-APR-1998; 98MO-US06956.

PR 10-APR-1997; 97US-0837312.

XX (GENW) GENNETICS INST INC.

PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racie LA, Spaulding V, Treacy M;

DR WPI: 1999-070078/06.

XX New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries

PS Claim 1, Page 399; 641pp; English.

XX The present sequence represents an expressed sequence tag (EST), and is
 CC a polynucleotide of the invention. The polynucleotides of the invention
 CC are all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.

XX Sequence 504 BP; 124 A; 160 C; 124 G; 96 T; 0 other;

Query Match 13.0%; Score 407.2; DB 20; Length 504;

Best Local Similarity 90.1%; Pred. No. 9.3e-94;

Matches 436; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

XX
 QY 1095 agtgaactcttgaagcacaagaagcagatctgcacagatagaatccatccaatt 1154
 DB 12 agagcctactctgaagcacaagaagatctgcacagatagaatccatccaatt 71
 QY 1155 ccacatctgcatcagcagcagcagcagcagcagcagcagcagcagcagcagc 1214
 DB 72 ccacatctgcatcagcagcagcagcagcagcagcagcagcagcagcagcagc 131
 QY 1215 tgcctccataatgacgctgtgctacacagcccaacagagccctgcagccactgta 1274
 DB 132 cgcctccataatgacgctgtgctacacagcccaacagagccctgcagccactgta 191
 QY 1275 ccaggtcgggtgcctctcagatccggtgttgcacaagcgggagacagctcag 1334
 DB 192 ccaggtcgggtgcctctcagatccggtgttgcacaagcgggagacagctcag 251
 QY 1335 gacatgtctcttatctgcacaacaaagacagctatgactatgtggtcagcag 1394
 DB 252 gacatgtctcttatctgcacaacaaagacagctatgactatgtggtcagcag 311
 QY 1395 ggcacgaagcgggtcccaagtcagatccctgtgatcagaagccctcttcagta 1454
 DB 312 ggcacgaagcgggtcccaagtcagatccctgtgatcagaagccctcttcagta 371
 QY 1455 caaggtgacacaccccatcaccacccacgctgcacatcacgctccctcggagaaat 1514
 DB 372 caaggtgacacaccccatcaccacccacgctgcacatcacgctccctcggagaaat 431
 QY 1515 gtggaaacaggaagacacctatactcagcagcggtgtgctgtggtgaatccac 1574
 DB 432 gtggaaacaggaagacacctatactcagcagcggtgtgctgtggtgaatccac 491
 QY 1575 tggc 1578
 DB 492 cacc 495
 RESULT 7
 ID 246041 standard; cDNA; 631 BP.
 AC 246041;
 XX 25-APR-2000 (first entry)
 DE Partial cDNA encoding a transferase designated HUTRAN-3 fragment.
 XX
 KW Transferase; HUTRAN-3; arginine methyltransferase; autoimmune disorder;
 KW inflammatory disorder; AIDS; atherosclerosis;
 KW adult respiratory distress syndrome; allergy; asthma; trauma;
 KW autoimmune thyroiditis; bronchitis; Crohn's disease; diabetes mellitus;
 KW gout; Grave's disease; osteoarthritis; osteoporosis; pancreatitis;
 KW psoriasis; rheumatoid arthritis; infection; neurological disorder;
 KW epilepsy; ischemic cerebrovascular disease; stroke; Alzheimer's disease;
 KW Picks disease; Huntington's disease; dementia; Parkinson's disease;
 KW extrapyramidal disorder; viral central nervous system disease;
 KW prion disease; central nervous system developmental disorder;
 KW neuroskeletal disorder; muscular dystrophy; neuromuscular disorder;
 KW peripheral nervous system disorder; mental disorder; schizophrenic;
 KW anxiety; reproductive disorder; gastrointestinal disorder;
 KW indigestion; gastritis; anorexia; nausea; abdominal angina;
 KW gastroenteritis; intestinal obstruction; peptic ulcer;
 KW irritable bowel syndrome; diarrhoea; constipation;
 KW gastrointestinal haemorrhage; cancer; ss.

OS Homo sapiens.
 XX
 XX WO200000594-A2.
 XX
 XX 06-JAN-2000.
 XX
 XX 29-JUN-1999; 99WO-US14651.
 XX
 XX 30-JUN-1998; 98US-0109204.
 XX
 XX (INCY-) INCYTE PHARM INC.
 XX
 XX Lal P, Pandman O, Hillman JL, Guegler KJ, Gorgone GA, Corley NC,
 XX Patterson C,
 XX WPI; 2000-147267/13.
 XX
 XX Novel human transferases used for the diagnosis, treatment, and
 XX prevention of autoimmune/inflammatory, neurological, reproductive and
 XX gastrointestinal disorders and cancer
 XX
 XX Disclosure: Page 89-90; 95pp; English.
 XX
 XX 246039-47 represent partial fragments which were used to produce Incyte
 XX clone 2525071, which encodes a transferase polypeptide, designated
 XX HUTRAN-3. The polypeptide is an arginine methyltransferase. The
 XX HUTRAN polypeptides, polynucleotides, agonists, antagonists, and
 XX antibodies can be used to diagnose, treat or prevent autoimmune/
 XX inflammatory diseases (e.g. AIDS, adult respiratory distress syndrome,
 XX allergies, asthma, atherosclerosis, autoimmune thyroiditis, bronchitis,
 XX Crohn's disease, diabetes mellitus, gout, Grave's disease,
 XX osteoarthritis, osteoporosis, pancreatitis, psoriasis, rheumatoid
 XX arthritis, infections, trauma, neurological disorders (e.g. epilepsy,
 XX ischemic cerebrovascular disease, stroke, Alzheimer's disease, Picks
 XX disease, Huntington's disease, dementia, Parkinson's disease and other
 XX extrapyramidal disorders, viral central nervous system disease, prion
 XX diseases, central nervous system developmental disorders, neuroskeletal
 XX disorders, muscular dystrophy, neuromuscular disorders, peripheral
 XX nervous system disorders, mental disorders, schizophrenic, anxiety,
 XX reproductive disorders and gastrointestinal disorders (e.g. dyspepsia,
 XX indigestion, gastritis, anorexia, nausea, abdominal angina,
 XX gastroenteritis, intestinal obstruction, intestinal tract infections,
 XX peptic ulcer, irritable bowel syndrome, diarrhoea, constipation,
 XX gastrointestinal haemorrhage, and cancer.
 XX
 XX Sequence 631 BP; 158 A; 157 C; 171 G; 129 T; 16 other;
 XX
 XX
 XX Query Match 9.1%; Score 284.2; DB 21; Length 631;
 XX Best Local Similarity 87.8%; Pred. No. 1.7e-62;
 XX Matches 332; Conservative 0; Mismatches 44; Indels 2; Gaps 2;
 QY 385 ttctacaatcctgaaaacctgtgcgggccaacactggagcgtctgttcaagtga 444
 DB 2 ttctacaatcctgaaaacctgtgcgggccaacactggagcgtctgttcaagtga 61
 QY 445 cggacagaggaactcctcagctgtgcagctactccagattctatgctacaccacagcag 504
 DB 62 cggacagaggaactcctcctcgtgcagctactccagattctatgctacaccacagcag 121
 QY 505 caggaatgcatcaggaactgtgcgagacgagcactaccacagcgtcgatccctgcagaaac 564
 DB 122 caggaatgcatcaggaactgtgcgagacgagcactaccacagcgtcgatccctgcagaaac 181
 QY 565 ccaacaggaactcaaggaacagatcgttctagatgtggtgtgtgctctggatcgtcga 624
 DB 182 ccaacaggaactcaaggaacagatcgttctagatgtggtgtgtgctctggatcgtcga 241
 QY 625 tttttgtcgtccaagcagagcaggaatattatgcaggtggaagcagcacaacat-ggc 683
 DB 242 tttttgtcgtccaagcaggtggaacaaattctacgcggtggaagcagcacaacat-ggc 301
 QY 684 tcagcatgagaggtctcgtgtggaagtaacaatctgac-agaccgcatcgtgtcatcc 742

XX (INCYTE) INCYTE PHARM INC.
PI Lal P, Bandman O, Hillman JL, Guegler KJ, Gorgone GA, Corley NC,
PI Patterson C;
XX WPI: 2000-147267/13.
XX
XX Novel human transferases used for the diagnosis, treatment, and
PT prevention of autoimmune/inflammatory, neurological, reproductive and
PT gastrointestinal disorders and cancer
XX
PS Disclosure: Page 89; 95pp; English.
XX
XX 246039-47 represent partial fragments which were used to produce Incyte
CC clone 2525071, which encodes a transferase polypeptide, designated
CC HGTBRAN-3. The polypeptide is an arginine methyltransferase. The
CC HMTBRAN polypeptides, polynucleotides, agonists, antagonists, and
CC antibodies can be used to diagnose, treat or prevent autoimmune/
CC inflammatory diseases (e.g. AIDS, adult respiratory distress syndrome,
CC Crohn's disease, diabetes mellitus, gout, Grave's disease,
CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, rheumatoid
CC arthritis, infections, trauma, neurological disorders (e.g. epilepsy,
CC ischemic cerebrovascular disease, stroke, Alzheimer's disease, Pick's
CC disease, Huntington's disease, dementia, Parkinson's disease and other
CC extrapyramidal disorders), viral central nervous system disease, prion
CC diseases, central nervous system developmental disorders, neuroskeletal
CC disorders, muscular dystrophy, neuromuscular disorders, peripheral
CC nervous system disorders, mental disorders, schizophrenic, anxiety,
CC reproductive disorders and gastrointestinal disorders (e.g. dyspepsia,
CC indigestion, gastritis, anorexia, nausea, abdominal angina,
CC gastroenteritis, intestinal obstruction, intestinal tract infections,
CC peptic ulcer, irritable bowel syndrome, diarrhoea, constipation,
CC gastrointestinal haemorrhage, and cancer.
XX
SQ Sequence 256 BP; 54 A; 77 C; 67 G; 58 T; 0 other;

Query Match 7.1%; Score 221.4; DB 21; Length 256;
Best Local Similarity 91.8%; Pred. No. 9.9e-47;
Matches 234; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 385 ttctacaacatctgaaacacctgtctggggcacaacactgagcgctctgttcagttag 444
DB 2 ttctacaacatctgaaacacctgtctggggcacaacactgagcgctctgttcagttag 61
OY 445 cggagagaggaatctctgctgtgtagtactctccagtttatgagctactatccagcag 504
DB 62 cggagagaggaatctctgctgtgtagtactctccagtttatgagctactatccagcag 121
OY 505 cagaacatgatgcagactatgtcgcagacagcactacacagctgcagtcctgcagac 564
DB 122 cagaacatgatgcagactatgtcgcagacagcactacacagctgcagtcctgcagac 181
OY 565 cacaacgagactcaaggaagaatcgttctagatgtgagctgtggtcttggatcctgtca 624
DB 182 cacaacgagactcaaggaagaatcgttctagatgtgagctgtggtcttggatcctgtcg 241
OY 625 tttttgtgtccaa 639
DB 242 tttttgtgtccaa 256

XX Human; colon cancer; tumour; diagnosis; gene expression product;
KW probe; detection; cancerous state; metastasis; identification;
KW breast cancer; oestrogen receptor-positive breast cancer; therapy;
KW oestrogen receptor-negative breast cancer; lung cancer; ss.
OS Homo sapiens.
XX
XX WO9558675-A2.
XX
XX 18-NOV-1999.
XX
XX
XX 13-MAY-1999; 99WO-US10602.
XX
XX 14-MAY-1998; 98US-0085426.
XX
XX 15-MAY-1998; 98US-0085537.
XX
XX 15-MAY-1998; 98US-0085696.
XX
XX 21-OCT-1998; 98US-0105234.
XX
XX 27-OCT-1998; 98US-0105877.
XX
XX (CHIR) CHIRON CORP.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
XX WPI: 2000-126369/11.
XX
XX Polynucleotide library used to determine cancerous states of mammalian
PT cells -
PT
PS Claim 1: Page 386; 1097pp; English.
XX
XX A00010 to A02716 represent polynucleotides isolated from cDNA libraries
CC constructed from human colon cancer cell lines. The present invention
CC also describes a method of detecting differentially expressed genes
CC correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived.
CC The polynucleotide sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of
CC premetastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer.
XX
SQ Sequence 300 BP; 67 A; 101 C; 74 G; 58 T; 0 other;

Query Match 6.5%; Score 203.6; DB 21; Length 300;
Best Local Similarity 85.6%; Pred. No. 3.5e-42;
Matches 238; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

OY 1375 atcagatgtgtgacagagtgagacagagctccacagctccagtaacctgtgactca 1434
DB 17 atcagatgtgtgacagagtgagacagagctccacagctccagtaacctgtgactcg 76
OY 1435 aagaacccctcttcaggtacacaggtacacacccatcacccacacacacacacac 1494
DB 77 aagaacccctcttcaggtacacaggtacacacccatcacccacacacacacacac 136
OY 1495 acgtctccctcggagatatgtggaacaggaagacacataatctcagcagcggtg 1554
DB 137 acatctccctcggagatatgtggaacaggaagacacacataatctcagcagcggtg 196
OY 1555 gcctgtgtgtgaatgtcctactgtcagacactgtgacagtggttatgtcgcggtccagt 1614

CC human protein. The gene numb

XX

AC Z46046;
XX
XX 25-APR-2000 (first entry)
DE Partial cDNA encoding a transferase designated HUTRAN-3 fragment.
XX
XX
XX Transferase: HUTRAN-3; arginine methyltransferase; autoimmune disorder;
XX inflammatory disorder; AIDS; atherosclerosis;
XX adult respiratory distress syndrome; allergy; asthma; trauma;
XX autoimmune thyroiditis; bronchitis; Crohn's disease; diabetes mellitus;
XX gout; Grave's disease; osteoarthritis; osteoporosis; pancreatitis;
XX psoriasis; rheumatoid arthritis; infection; neurological disorder;
XX epilepsy; ischemic cerebrovascular disease; stroke; Alzheimer's disease
XX Picks disease; Huntington's disease; dementia; Parkinson's disease;
XX extrapyramidal disorder; viral central nervous system disease;
XX prion disease; central nervous system developmental disorder;
XX neuroskeletal disorder; muscular dystrophy; neuromuscular disorder;
XX peripheral nervous system disorder; mental disorder; schizophrenic;
XX anxiety; reproductive disorder; gastrointestinal disorder; dyspepsia;
XX indigestion; gastritis; anorexia; nausea; abdominal angina;
XX gastroenteritis; intestinal obstruction; peptic ulcer;
XX irritable bowel syndrome; diarrhoea; constipation;
XX gastrointestinal haemorrhage; cancer; ss.
XX
XX Homo sapiens.
XX
XX
XX WO200000594-A2.
XX
XX 06-JAN-2000.
XX
XX 29-JUN-1999; 99WO-US14651.
XX
XX 30-JUN-1998; 98US-0109204.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Lai P, Bandman O, Hillman JL, Guegler KJ, Gorgone GA, Corley NC;
XX Pi Patterson C;
XX WPI: 2000-147267/13.
XX
XX Novel human transferases used for the diagnosis, treatment, and
XX prevention of autoimmune/inflammatory, neurological, reproductive and
XX gastrointestinal disorders and cancer
XX
XX Disclosure: Page 92; 95pp; English.
XX
XX Z46039-47 represent partial fragments which were used to produce Incyte
XX clone 2525071, which encodes a transferase polypeptide, designated
XX HUTRAN-3. The polypeptide is an arginine methyltransferase. The
XX HUTRAN polypeptides, polynucleotides, agonists, antagonists, and
XX antibodies can be used to diagnose, treat or prevent autoimmune/
XX inflammatory diseases (e.g. AIDS, adult respiratory distress syndrome,
XX allergies, asthma, atherosclerosis, autoimmune thyroiditis, bronchitis,
XX Crohn's disease, diabetes mellitus, gout, Grave's disease,
XX osteoarthritis, osteoporosis, pancreatitis, psoriasis, rheumatoid
XX arthritis, infections, trauma, neurological disorders (e.g. epilepsy,
XX ischemic cerebrovascular disease, stroke, Alzheimer's disease, Picks
XX disease, Huntington's disease, dementia, Parkinson's disease and other
XX extrapyramidal disorders), viral central nervous system disease, prion
XX diseases, central nervous system developmental disorders, neuroskeletal
XX disorders, muscular dystrophy, neuromuscular disorders, peripheral
XX nervous system disorders, mental disorders, schizophrenic, anxiety,
XX reproductive disorders and gastrointestinal disorders (e.g. dyspepsia,
XX indigestion, gastritis, anorexia, nausea, abdominal angina,
XX gastroenteritis, intestinal obstruction, intestinal tract infections,
XX peptic ulcer, irritable bowel syndrome, diarrhoea, constipation,
XX gastrointestinal haemorrhage, and cancer.
XX
XX Sequence 529 BP; 98 A; 154 C; 139 G; 111 T; 27 other;

Best Local Similarity	77.0%;	Pred: No. 6.9e-26;
Matches	218;	Conservative
	0;	Mismatches
	55;	Indels
	10;	Gaps
	4;	

QY	2295	gctcagaccctcctctcaatcacacgagccacatgctacatggtggcgtgacatgctctt	2354
Db	34	gcgagggcctccctcttcgaagacagagcctcc--gttcaacaagagtgatgaatgcgtctt	91
QY	2355	ttttaattattatttttttacgaaagaacacagtgctaacccacagacccctctgagaac-	2414
Db	92	ttttaattattattttttttttagaaagaacacagtgctaacctcgcacacccctctgtaaac	151
QY	2415	cgcgtg-cgcgcgaagccgacgagcccccgttcctcagtcgagcgaag-tttcagtgtgaag	2472
Db	152	agggccgcgcggcgagccagcagcggccctctccctagatcaggaagcgcgcgggggaagg	211
QY	2473	gtgagccctgtcaagccttcagagtggtgcagagccctccacacaaagggttcaactcaa	2533
Db	212	gtmcccgcgcgaggtctcagg-----nnmccctccacacaaagggttcaactcaca	265
QY	2533	cttgaatgtgtcaaacacacccagtggtcgaaggctagtcctt	2575
Db	266	cttgaaatgtgtcaaacacacccagtggtcgaaggctcgttcct	308

RESULT	13
Q61313	
ID	Q61313 standard; DNA; 217 BP.

AC Q61313;

DT 16-MAR-1994 (first entry)

Human brain Expressed Sequence Tag EST01321.

KM Gene transcription product; genetic markers; tagging; in vivo;
transcription; mapping; locations; chromosomes; chromosomal, ss.
KM

OS Homo sapiens

PN W09316178-A.

PD 19-AUG-1993.

PF 12-FEB-1993; 93WO-US01321.

PR 12-FEB-1992; 92US-0837195.

PA (USSH) US DEPT HEALTH & HUMAN SERVICE.

PI Adams MD, Moreno RF, Venter CJ;

DR WPI; 1993-272882/34.

PT Enriched oligonucleotides and corresp. sequences - used as

PT of most human genes

PS Example 4; Page 471; 500pp; English.

The expressed sequence tag was isolated from a human brain cDNA library as part of a large set of ESTs which can be used as markers for human genes transcribed *in vivo*. They can be used to facilitate tagging of most human genes; for mapping locations of expressed genes on chromosomes; for individual or forensic identification; for mapping locations of disease-associated genes; for identification of tissue type, and for prep. of antisense sequences, probes and constructs. EST01321 has an "excellent" coding probability as evaluated using the coding-region prediction program CRM. See also Q59041-Q61440.

SQ Sequence 217 BP; 45 A; 76 C; 61 G; 30 T; 5 other,

Query Match	4.58;	Score 139.8;	DB 21;	Length 529;
-------------	-------	--------------	--------	-------------

```
Query match      3.98;  Score 123.2;  DB 14;  Length 217;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 09:55:06 ; Search time 23.09 Seconds
(without alignments)
1505.208 Million cell updates/sec

Title: US-09-464-377-2

Perfect score: 3168
Sequence: 1 MAANAATVAGPAGSAGVAG.....PAISNAPSISPTMTMYGS 608

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.0401.*
1: /SID56/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2826.5	89.2	551	21	B43316 Human ORFX ORF3080
2	2290.5	72.3	447	21	Y54593 Human acid sequence
3	498.5	15.7	390	21	G24704 Arabidopsis thalia
4	498	15.7	341	21	G24705 Arabidopsis thalia
5	494.5	15.6	353	19	W64203 Arginine methyltra
6	488.5	15.4	357	21	B43525 Human cancer assoc
7	488.5	15.4	361	19	W64202 Human interferon r
8	474.5	15.0	531	21	Y84432 Amino acid sequenc
9	472.5	14.9	306	21	G24706 Arabidopsis thalia
10	467.5	14.8	339	21	G29547 Arabidopsis thalia
11	467.5	14.8	366	21	G29546 Arabidopsis thalia

12	443.5	14.0	306	21	G29548
13	378	11.9	360	19	W64204
14	299.5	9.5	297	20	Y48597
15	267	8.4	122	20	Y26923
16	183	5.8	97	21	G01996
17	182.5	5.8	193	21	B42019
18	177	5.6	692	21	B49313
19	169.5	5.4	589	19	W54096
20	165	5.2	542	21	G21384
21	165	5.2	599	21	G21383
22	165	5.2	645	21	G21382
23	127.5	4.0	276	19	W30525
24	117	3.7	531	19	W56311
25	114	3.6	1275	20	Y29084
26	114	3.6	1275	21	Y44262
27	108	3.4	356	14	R37715
28	107.5	3.4	1289	13	R20067
29	107.5	3.4	1289	13	R28890
30	107.5	3.4	1289	13	R29027
31	107.5	3.4	1289	13	R28810
32	107.5	3.4	1289	13	R29517
33	107.5	3.4	1289	14	R44202
34	107.5	3.4	1289	15	R58632
35	107.5	3.4	1289	16	R76113
36	107.5	3.4	1289	16	W13885
37	107.5	3.4	1289	21	B13892
38	99	3.1	552	12	R13879
39	98.5	3.1	1071	21	Y97317
40	98.5	3.1	3782	21	Y77179
41	98	3.1	235	19	W98690
42	98	3.1	777	20	Y43091
43	97.5	3.1	331	21	G04547
44	97.5	3.1	331	21	G43327
45	97.5	3.1	985	18	W15191

ALIGNMENTS

RESULT	1	
B43316	08-FEB-2001 (first entry)	
ID	B43316 standard; Protein; 551 AA.	
AC	B43316;	
DT	08-FEB-2001 (first entry)	
XX	Human ORFX ORF3080 polypeptide sequence SEQ ID NO:6160.	
XX	Human: Open reading frame; ORFX; detection; cytosolic; hepatotropic;	
KW	vulnerable; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;	
KW	anticonvulsant; osteoporotic; antirheumatic; immunosuppressant; cardiant;	
KW	immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;	
KW	hypotensive; dermatological; immunosuppressive; antineoplastic;	
KW	antiviral; antibacterial; antifungal; antineumatic; antithyroid;	
KW	antianaemic; gene therapy; cancer; proliferative disorder; hypertension;	
KW	neurodegenerative disorder; osteoarthritis; graft vs host disease;	
KW	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;	
KW	cholesterol ester storage; systemic lupus erythematosus; infection;	
KW	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;	
KW	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;	
KW	bone damage; cartilage damage; antiinflammatory disease; coagulation;	
KW	thrombosis; contraceptive.	
XX		
OS	Homo sapiens.	
XX		
PN	W0200058473-A2.	
XX		
PD	05-OCT-2000.	
XX		
PF	31-MAR-2000; 2000MO-US08621.	
XX		
PR	31-MAR-1999; 99US-0127607.	

Arabidopsis thalia
Human HCP-1. Homo
Human breast tumor
Amino acids 58-179
Human secreted pro
Human ORFX ORF1783
Mouse brown adipos
Homo sapiens TGL16
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Kurthia sp. bioc g
Protein disulphide
Human C1F130 prote
Human cell cycle r
Carnitine 4-O-m
B.thuringiensis to
Toxin 17b. Bacill
Bacillus thuringie
Bt toxin 17b. Bac
Bt toxin 17b. Bac
Bacillus thuringie
Bacillus thuringie
Ps17b acaride-acti
17b toxin. Bacill
Bacillus thuringie
Indoleacetic acid
S.vadenosylmethoni
S. venezuelae deso
H. pylori GHP 611
Mouse semaphorin H
Arabidopsis thalia
Arabidopsis thalia
Aspergillus oryzae

PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127228.
 PR 30-MAR-2000; 2000US-0540763.
 XX (CURA-) CURAGEN CORP.
 PA Shinkets RA, Leach M;
 PI WPI, 2000-602362/57.
 DR N-PSDB; C77525.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 PS
 PS Claim 11; Page 5346-5348; 5507pp; English.
 XX
 CC C7446 to C77606 encode the proteins given in B40237 to B43397, which
 CC represent the human ORF open reading frames 1 to 3161. The ORF
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiproliferative; antiparinsonian; nocitropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiatic; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antihemmatic; antihypertoid; and antianemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORF-associated disorder. The nucleic acids can be used to express ORF
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 XX Sequence 551 AA:
 SQ

Query Match 89.2%; Score 2826.5; DB 21; Length 551;
 Best Local Similarity 98.4%; Pred. No. 2.7e-234;
 Matches 542; Conservative 6; Mismatches 2; Indels 1; Gaps 1;

QY 59 LEVRAGPAAAGIALYSHEDVCFKCSVSRETECSRVRGROSFITTLGCNSVLIQFAPPHDF 118
 DB 1 LEVRAGPAAAGIALYSHEDVCFKCSVSRETECSRVRGROSFITTLGCNSVLIQFAPPHDF 60
 QY 119 CSFYNIKTRCHRTLEKRSYFSESTRESSAVOYFOFGYISQOONMADYVRCGYORAIL 178
 DB 61 CSFYNIKTRCHRTLEKRSYFSESTRESSAVOYFOFGYISQOONMADYVRCGYORAIL 120
 QY 179 QHRTPEKDKIVADVCGSGSIIISFFAQAQARKIYAEASTMAQHAELVKNLTPRIYV 238
 DB 121 QHRTPEKDKIVADVCGSGSIIISFFAQAQARKIYAEASTMAQHAELVKNLTPRIYV 180
 QY 239 IPGKVEEVSLEPQVDIIISEPMGYMLFNERMLESYLHAKKYLKPSGNMPEPTIGDVLAPF 298
 DB 181 IPGKVEEVSLEPQVDIIISEPMGYMLFNERMLESYLHAKKYLKPSGNMPEPTIGDVLAPF 240
 QY 299 TDEOLYMEQFTANRKYOSFHFVNDLSALRGAAVDEYFQPVVDPTDIRLMAKSKYRV 358
 DB 241 TDEOLYMEQFTANRKYOSFHFVNDLSALRGAAVDEYFQPVVDPTDIRLMAKSKYRV 300
 QY 359 NPLEAKGGLHRIEIPFKPHMLHSGLVHGLAFWEDVAFGSIPTVMLESTAPTEPLTHWYO 418
 DB 301 NPLEAKGGLHRIEIPFKPHMLHSGLVHGLAFWEDVAFGSIPTVMLESTAPTEPLTHWYO 360
 QY 419 VRCLEQSPLEFARAGDTLSCTCLLIANKROSYDISIVAQVDGTSKSSNLLDKNPFERYT 478
 DB 361 VRCLEQSPLEFARAGDTLSCTCLLIANKROSYDISIVAQVDGTSKSSNLLDKNPFERYT 420

QY 479 GTTPSPPGSHYTSPEENMNTGSTYNLSSGVAVAGMPTAYDLDSVIAGSSVGHNNLIP 538
 DB 421 GTTPSPPGSHYTSPEENMNTGSTYNLSSGVAVAGMPTAYDLDSVIAGSSVGHNNLIP 480
 QY 539 LANTGIVNHTSRMGSSIMSTGIVGSSGAQG-GGGSSSAHYAVNNOFTMGGAISMASPM 597
 DB 481 LANTGIVNHTSRMGSSIMSTGIVGSSGAQG-GGGSSSAHYAVNNOFTMGGAISMASPM 540
 QY 598 SIPTNTHYGS 608
 DB 541 SIPTNTHYGS 551
 RESULT 2
 Y54593
 ID Y54593 standard; Protein; 447 AA.
 XX
 AC Y54593;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Amino acid sequence of a human transferrase designated HUFAN-3.
 XX
 KW Transferrase; HUFAN-3; arginine methyltransferase; autoimmune disorder;
 KW inflammatory disorder; AIDS; atherosclerosis;
 KW adult respiratory distress syndrome; allergy; asthma; trauma;
 KW autoimmune thyroiditis; bronchitis; Crohn's disease; diabetes mellitus;
 KW gout; Grave's disease; osteoarthritis; osteoporosis; pancreatitis;
 KW psoriasis; rheumatoid arthritis; infection; neurological disorder;
 KW epilepsy; ischemic cerebrovascular disease; stroke; Alzheimer's disease;
 KW picks disease; Huntington's disease; dementia; Parkinson's disease;
 KW extrapyramidal disorder; viral central nervous system disease;
 KW prion disease; central nervous system developmental disorder;
 KW neuroskeletal disorder; muscular dystrophy; neuromuscular disorder;
 KW peripheral nervous system disorder; mental disorder; schizophrenic;
 KW anxiety; reproductive disorder; gastrointestinal disorder; dyspepsia;
 KW indigestion; gastritis; anorexia; nausea; abdominal angina;
 KW gastroenteritis; intestinal obstruction; peptic ulcer;
 KW irritable bowel syndrome; diarrhoea; constipation;
 KW gastrointestinal haemorrhage; cancer.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..46
 FT Modified-site /note= "signal sequence"
 FT 18
 FT Modified-site /note= "potential glycosylation site"
 FT 27..35
 FT Region /note= "methyltransferase motif"
 FT 49..53
 FT Modified-site /note= "methyltransferase motif"
 FT 69
 FT Modified-site /note= "potential glycosylation site"
 FT 71
 FT Modified-site /note= "potential phosphorylation site"
 FT 85
 FT Modified-site /note= "potential phosphorylation site"
 FT 88..95
 FT Region /note= "methyltransferase motif"
 FT 117..126
 FT Modified-site /note= "methyltransferase motif"
 FT 127
 FT Modified-site /note= "potential phosphorylation site"
 FT 165..169
 FT Region /note= "methyltransferase motif"
 FT 191
 FT Modified-site /note= "potential phosphorylation site"
 FT 286
 FT Modified-site /note= "potential phosphorylation site"
 FT 293
 FT Misc-difference /note= "potential phosphorylation site"
 FT /note= "ala encoded by CC"
 FT Modified-site 330

FT Modified-site /note= "potential phosphorylation site"
 FT 343
 FT Modified-site /note= "potential glycosylation site"
 FT 355
 FT Modified-site /note= "potential phosphorylation site"
 FT 384
 FT Modified-site /note= "potential glycosylation site"
 PN WC200000594-A2.
 XX 06-JAN-2000.
 XX 29-JUN-1999; 99WO-US14651.
 XX 30-JUN-1998; 98US-0109204.
 XX (INCY-) INCYTE PHARM INC.
 PA Lal P, Bandman O, Hillman JL, Guegler KU, Gorgone GA, Corley NC,
 PI Patterson C;
 DR WPI: 2000-147267/13.
 DR N-PSDB; 246024.
 XX
 PT Novel human transferases used for the diagnosis, treatment, and
 PT prevention of autoimmune/inflammatory, neurological, reproductive and
 PT gastrointestinal disorders and cancer -
 XX
 PS Claim 1; Page 79-80; 95pp; English.
 XX
 CC The present sequence represents a transferase polypeptide, designated
 CC HTRAN-3. The HTRAN cDNA sequence is derived from Incyte clone
 CC 255071. The sequence is obtained from the partial sequences 246039-47.
 CC The polypeptide is a arginine methyltransferase. The HTRAN polypeptides,
 CC polynucleotides, agonists, antagonists, and antibodies can be used to
 CC diagnose, treat or prevent autoimmune/inflammatory diseases (e.g. AIDS,
 CC adult respiratory distress syndrome, allergies, asthma, atherosclerosis,
 CC autoimmune thyroiditis, bronchitis, Crohn's disease, diabetes mellitus,
 CC gout, Grave's disease, osteoarthritis, osteoporosis, pancreatitis,
 CC psoriasis, rheumatoid arthritis, infections, trauma, neurological
 CC disorders (e.g. epilepsy, ischemic cerebrovascular disease, stroke,
 CC Alzheimer's disease, Pick's disease, Huntington's disease, dementia,
 CC Parkinson's disease and other extrapyramidal disorders), viral central
 CC nervous system disease, prion diseases, central nervous system
 CC developmental disorders, neuroskeletal disorders, muscular dystrophy,
 CC neuromuscular disorders, peripheral nervous system disorders, mental
 CC disorders, schizophrenia, anxiety, reproductive disorders and
 CC gastrointestinal disorders (e.g. dyspepsia, indigestion, gastritis,
 CC anorexia, nausea, abdominal angina, gastroenteritis, intestinal
 CC obstruction, intestinal tract infections, peptic ulcer, irritable bowel
 CC syndrome, diarrhoea, constipation, gastrointestinal haemorrhage, and
 CC cancer.
 CC
 XX
 XX Sequence 447 AA;
 SQ

Query Match 72.3%; Score 2290.5; DB 21; Length 447;
 Best Local Similarity 96.7%; Pred. No. 2.2e-186;
 Matches 441; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 163 MNOQYVTRGTQVRAIIHQHTDFKDKIVLDVGGSGIISFPAQAQARKIYAEASTMOH 222
 Db 1 mmqdyvrtgtygralqhtdfkdkivldvgsgsllsffaagagarklyaeastmahn 60
 QY 223 AEVLKSNLDRIVIVIGKVEVSLPQVDIIISEPMGYMFMFNMLESYLNARKYIKP 282
 Db 61 aeavlksnldrlivivlgkveevslpgevdlisepmgymfmfnmlesylhakkylkp 120
 QY 283 SGNNEPTIGDVHLAPFDDEQLMEQFTANRKYQPSFHGVDSALRGAANDERYQOPVVD 342
 Db 121 sgnmfptgldvhlapfddeqlymeqftkanfkyqpsfhgvdsalrgaavdeyfrqpvd 180
 QY 343 TFDRIILMAKSVKTYVNLFAKEGDLHRIEIPKFNHLSGVLHGLAMFDVAFGSINT 402

Db 181 tldirilmaksvkytwnflaekegdhlrlieipkftlmhleglvhglafvdfvfigsimt 240
 QY 403 VMLSTAPTEPLFMWVROCLFOSPLFAKAGDPTSGICLLIANKROSVDISIVQOFTGS 462
 Db 241 vmlstapteplfmwvrcldfospflakagdtlsgiclliankrqsdlsivagvqftgs 300
 QY 463 KSNMLDLKNPFRYTGTPSPPGSHYTPSENMMNTGSTYMLSSGAVAGMPTAYDLS 522
 Db 301 ksnmldlnkpnfrtygtspppgshytpsenmmntgstymlssgavagmptaydls 360
 QY 523 SVIAGSSVGHNNLPLANTGIVNHTSRMGSTMGSTGYOGSSGAGC-GGSSSAHYAVN 581
 Db 361 sviasgssvghnnlplantgivnhtsrmsgsmstgtyvssgagqsgsgstcsahyavn 420
 QY 582 NOFTMGCPAISMASPMKSIPTNTHYGS 608
 Db 421 sqftmgcpaismaspmksiptnthys 447

RESULT 3
 G24704
 ID G24704 standard; Protein: 390 AA.
 XX
 AC G24704;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 28481.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0128845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139617.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144684.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145313.
PR 27-JUL-1999: 99US-0145918.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.

PR 02-AUG-1999: 99US-0146389.
PR 03-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 04-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.
PR 12-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148565.
PR 13-AUG-1999: 99US-0148684.
PR 16-AUG-1999: 99US-0149368.
PR 17-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149929.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 08-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 18-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.

	RESULT	7
ID	M64202	
XX	M64202 standard; Protein; 361 AA.	
AC		
XX	M64202;	
DT	09-NOV-1998 (first entry)	
DE	Human interferon receptor 1 binding protein IRIB4.	
XX		
KW	Interferon receptor 1 binding protein; IRIB4; human; tumour; cancer; gene therapy; tissue graft; graft survival; arginine methyltransferase.	
KM		
RN	Homo sapiens.	
OS		
PN	W098J31796-A1.	
PD	23-JUL-1998.	
PF	15-JAN-1998; 98WO-US00671.	
PR	15-JAN-1997; 97US-0035636.	
PS	(MCIN/) MCINNIS P A. (YEDA) YEDA RES & DEV CO LTD.	
PI	Abramovitch C, Chebach JE, Revel M;	
DR	WI: 1998-414096/35.	
DR	N-PSDB; V44275.	
PT	New isolated interferon receptor binding proteins - used to develop products for modulating sensitivity to interferon, e.g. in the treatment of tumours or for prolonging graft survival	
PS	Claim 1; Page 39-40; 64pp; English.	
XX	This is a novel human protein, designated interferon receptor binding protein 4 (IRIB4), which interacts with the intracytoplasmic (IC) domain of the IFNAR1 chain of the interferon type 1 (IFN-alpha, beta or omega) receptor. IRIB4 has sequence homology to enzymes which utilize S-adenosyl methionine for methylating arginine residues in proteins, such as the arginine methyltransferase PRMT1 (see W64203). It was identified in a two-hybrid screening for proteins interacting with the IFNAR1-IC domain; another protein, IRIB1 (see W64199), was similarly identified. A cDNA clone (see W44275) encoding IRIB4, host cells and expression vectors are claimed. DNA encoding IRIB1 and IRIB4 can be used in cancer therapy where the increased cellular response to IFN would result in a decrease in malignant cell growth and an enhanced response to exogenous IFN therapy. Antisense IRIB1 or IRIB4 nucleic acids can be used for prolonging tissue or organ graft survival in patients as the rejection of these grafts in the host is mediated by the histocompatibility antigens (MHC class I) whose synthesis depends on the IFN stimulus. The products can also be used in detection and diagnosis.	
SQ	Sequence 361 AA;	
Query Match	15.4%; Score 488.5; DB 19; Length 361;	
Best Local Similarity	32.1%; Pred. No. le-33;	
Matches 120; Conservative	68; Mismatches 149; Indels 37; Gaps	
OY	97 OSFITTLCNSTLLOFPATPHDFCSFYNLTKRCRTHRSVSFRTERESSAVOYFOFGY 156	
Db	2 enfvati-ngamslqprleevscg-----gseeskpnaedmtskdyfyday 49	
OY	157 LSGCMMADVDYRTGYRALIHTDPKKIVLDVGCGSSILSFPAAGAKRIIVEA 216	

Db	50	fglnemlkdevrtlltyrnmfhnrlfkdkvvlwvsgsgllcmfaakgarkvlgiee	109
Qy	217	STMAQHAELVLYKSNNTFDRIVLPGRKEEYSLP-EGVDYIIISBPGMYLNFENRMLESYLH	275
Db	110	ssisdgavkikanklkhvvtllkxgveevelpjevkdilissemgyclfyesmlntvly	169
Qy	276	AK-KYLWPSGMPEPTIDVHLAPRTDQGLVMEQTKANFPYQPSFHHVDLSALRGAAVDE	334
Db	170	ardkwlpdpdglilfpdratllyvtaded-----rgykdyklmwenvygfamcslkdvali--	222
Qy	335	YFROPVVDTEPDIRILMKSVKTYVNFPEAKEGDILHREI-----PFKPHMLHSGLVHG	387
Db	223	--keplrdvvpkqlv-----tnaclikevdltytkvedlftspfcqlvkrndyvha	273
Qy	368	LAEPEDVAFTSISMTVWLSTAPTEPLTHWYQVRCFLQSPLEFARAGDPLSGTCLLIANKRQ	447
Db	274	lvayfnlefrchkrtrftgstespespylhwkxvfyfymedyltvkrgeelfgtlgmprnkn	333
Qy	448	SYDSIYAGVNDQG 461	
Db	334	nrldlftldldfkq 347	
RESULT 8			
ID	Y84432	Y84432 standard; Protein; 531 AA.	
XX	Y84432;		
DT	25-JUL-2000	(first entry)	
DE	Amino acid sequence of a human RNA-associated protein.		
XX			
KW	Human; RNA-associated protein; cell proliferation; cancer; inflammation;		
KW	immune response; reproductive disorder; actinic keratosis;		
KW	atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;		
KW	mixed connective tissue disease; myelofibrosis; primary thrombocythemia;		
KW	paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;		
XX	trauma.		
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	Modified-site	27	/note= "potential phosphorylation site"
FT	Modified-site	58	/note= "potential phosphorylation site"
FT	Modified-site	59	/note= "potential phosphorylation site"
FT	Modified-site	89	/note= "potential phosphorylation site"
FT	Modified-site	155	/note= "potential phosphorylation site"
FT	Modified-site	157	/note= "potential glycosylation site"
FT	Modified-site	242	/note= "potential phosphorylation site"
FT	Modified-site	339	/note= "potential phosphorylation site"
FT	Modified-site	428	/note= "potential phosphorylation site"
FT	Modified-site	430	/note= "potential phosphorylation site"
FT	Modified-site	439	/note= "potential phosphorylation site"
FT	Modified-site	475	/note= "potential phosphorylation site"
FT	Modified-site	492	/note= "potential phosphorylation site"
FT	Modified-site	522	/note= "potential phosphorylation site"
FT	Modified-site	523	/note= "potential glycosylation site"
FT	Modified-site		/note= "potential glycosylation site"

PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139765.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139889.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141847.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	20-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	03-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148365.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155139.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT	

OY 341 VDTDIRILMAKS-VKIVVNFLEAKEGDILHRIEIPFKFHMHSLGVHGLAFNFDVAFLGS 399
||| : : : : :
Db 172 vdtvdqngivdrslrllkmdiskmessga-sftapfkIvagnDylhalvayfdvstfmc 230
OY 400 IMTWLSTAPPEPLTHWQVRCLEFSPLEFAKAGDILSGCLLIANKROSYDLSI 453
||| : : : : :
Db 231 hklIgfstcpkrachmkqvtlyledvltlcegetltgmsvsnkknpridli 284
RESULT 10
G29547
ID G29547 standard; Protein; 339 AA.
XX
AC G29547;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35175.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137232.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.

PR	10-AUG-1999;	99US-0148319;
PR	11-AUG-1999;	99US-0148319;
PR	12-AUG-1999;	99US-0148341;
PR	13-AUG-1999;	99US-0148565;
PR	13-AUG-1999;	99US-0148684;
PR	16-AUG-1999;	99US-0149368;
PR	17-AUG-1999;	99US-0149175;
PR	18-AUG-1999;	99US-0149426;
PR	20-AUG-1999;	99US-0149722;
PR	20-AUG-1999;	99US-0149723;
PR	20-AUG-1999;	99US-0149929;
PR	23-AUG-1999;	99US-0149929;
PR	23-AUG-1999;	99US-0149930;
PR	25-AUG-1999;	99US-0150566;
PR	26-AUG-1999;	99US-0150884;
PR	27-AUG-1999;	99US-0151065;
PR	27-AUG-1999;	99US-0151066;
PR	27-AUG-1999;	99US-0151080;
PR	30-AUG-1999;	99US-0151303;
PR	31-AUG-1999;	99US-0151438;
PR	01-SEP-1999;	99US-0151930;
PR	07-SEP-1999;	99US-0152363;
PR	10-SEP-1999;	99US-0153070;
PR	13-SEP-1999;	99US-0153758;
PR	15-SEP-1999;	99US-0154039;
PR	16-SEP-1999;	99US-0154779;
PR	20-SEP-1999;	99US-0154779;
PR	22-SEP-1999;	99US-0155139;
PR	23-SEP-1999;	99US-0155486;
PR	24-SEP-1999;	99US-0155659;
PR	28-SEP-1999;	99US-0156458;
PR	29-SEP-1999;	99US-0156596;
PR	04-OCT-1999;	99US-0157117;
PR	05-OCT-1999;	99US-0157753;
PR	06-OCT-1999;	99US-0157865;
PR	07-OCT-1999;	99US-0158029;
PR	08-OCT-1999;	99US-0158232;
PR	12-OCT-1999;	99US-0158369;
PR	13-OCT-1999;	99US-0159293;
PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159295;
PR	14-OCT-1999;	99US-0159330;
PR	14-OCT-1999;	99US-0159331;
PR	14-OCT-1999;	99US-0159637;
PR	14-OCT-1999;	99US-0159638;
PR	18-OCT-1999;	99US-0159584;
PR	21-OCT-1999;	99US-0160741;
PR	21-OCT-1999;	99US-0160767;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160770;
PR	21-OCT-1999;	99US-0160814;
PR	21-OCT-1999;	99US-0160815;
PR	22-OCT-1999;	99US-0160980;
PR	22-OCT-1999;	99US-0160981;
PR	22-OCT-1999;	99US-0160989;
PR	22-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161920;
PR	28-OCT-1999;	99US-0161922;
PR	28-OCT-1999;	99US-0161993;
PR	29-OCT-1999;	99US-0162142;

Query Match	14.8%	Score 467.5	DB 21	Length 339
Best Local Similarity	35.9%	Pred. No. 6e-33		
Matches 112: Conservative	54	Mismatches 133	Indels 13	Gaps 6
QY	145	SSAVQFQFYGLSQOQNNMODYVRTGTGYORALLNHTDFKDKIKYLDVCGSGGIIISFFAA	204	

Accession	Gene	Protein	Function
Db 16	tsaaydysysfhigtheemlkdvrttsygdvlykknfllkdkvlidvsgtglalslca	75	
Qy 205	QAGARKITAVEASTMAQHAELVYKSNLTDRIIVIPGKVEEVSIP-EQVDIIITSEPGIM	263	
Db 76	kagahayavecsqmadtakeivksngfsdvltylkjleelipvbkdvdlissewmyf	135	
Qy 264	LFNEMLESTYIMAR-KYLKPSGNMFPIGTGVHLAPFDDEOLYMEQFKANFRQPSHGV	322	
Db 136	llyemltdvlyanrkvlvdgvlvlpdkaslylvalaed----ahykddkvetlwdvdygt	190	
Qy 323	DLSLRGAADVEYRROPVDFD-IRILMAKSVKYTVNPLEAKEGDLHRIEIPFKFHLH	381	
Db 191	dmscikrraice----plvdtvdngnlvtcdsklltkmdiskmaagda-sftapfklyagtr	245	
Qy 382	SGLVHGLAFWEDVAFIGSIMTWLSTAPTEPLTHMYQVRCILFQSPFLAKAGDTLSTCLL	441	
Db 246	ndhhalvaydvsftmchckmfgstgksrathwqtlvlyledvltlcegetltsmtl	305	
Qy 442	IANKRQSTDISI 453		
Db 306	aqnkknpdvqvl 317		
RESULT 11			
G29546			
ID 1	G29546 standard; Protein; 366 AA.		
XX AC	G29546;		
XX DE	17-OCT-2000 (first entry)		
DE DE	Arabidopsis thaliana protein fragment SEQ ID NO: 35174.		
XX KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.		
XX OS	Arabidopsis thaliana.		
XX PN	EP1033405-A2.		
PD PD	06-SEP-2000.		
XX PF	25-FEB-2000; 2000EP-0301439.		
XX PR	25-FEB-1999; 99US-0121825.		
PR 05-MAR-1999;	99US-0123180.		
PR 09-MAR-1999;	99US-0123548.		
PR 23-MAR-1999;	99US-0125788.		
PR 25-MAR-1999;	99US-0126264.		
PR 29-MAR-1999;	99US-0126785.		
PR 01-APR-1999;	99US-0127462.		
PR 06-APR-1999;	99US-0128234.		
PR 08-APR-1999;	99US-0128714.		
PR 16-APR-1999;	99US-0129845.		
PR 19-APR-1999;	99US-0130077.		
PR 21-APR-1999;	99US-0130449.		
PR 23-APR-1999;	99US-0130510.		
PR 23-APR-1999;	99US-0130891.		
PR 28-APR-1999;	99US-0131449.		
PR 30-APR-1999;	99US-0132048.		
PR 30-APR-1999;	99US-0132407.		
PR 04-MAY-1999;	99US-0132484.		
PR 05-MAY-1999;	99US-0132485.		
PR 06-MAY-1999;	99US-0132486.		
PR 06-MAY-1999;	99US-0132487.		
PR 07-MAY-1999;	99US-0132863.		
PR 11-MAY-1999;	99US-0134256.		
PR 14-MAY-1999;	99US-0134218.		
PR 14-MAY-1999;	99US-0134219.		
PR 14-MAY-1999;	99US-0134221.		
PR 14-MAY-1999;	99US-0134370.		
PR 18-MAY-1999;	99US-0134768.		

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 14.8%; Score 467.5; DB 21; Length 366;
Best Local Similarity 35.9%; Pred. No. 6; Be-32;
Matches 112; Conservative 54; Mismatches 133; Indels 13; Gaps 6;
OY 145 SSAVVOFOFGVYQQOONMADYVGTGYORALIQNTDFDKTVLVVCGSGILSEFAA 204
Db 43 tsadyyfidsyshfghieemikdvvtksyqdvlyknkflkdkklydvagagylslsica 102
OY 205 QAGARKIYAEASTMACHAELVYKSNMLTDRIYVPGKVEEVSIP-EQVDIIISEPMGYM 263
Db 103 kagaahyavescsmadtakelvyksngfsdvltylkjgkielelpyvpkvdvllsemgylf 162
OY 264 LFMERMLESYLHAK-KYLKSGNMFPRTIGVHLAPFDQOLYMEQFTKANRYPSPFHGV 322
Db 163 llyemldtvllyarkvlgvglylvpkkslyvtaled----ahykdxkvefddvygf 217
OY 323 DLGALRGAAVDEYFRQPVDTFD-IRILMAKSVKYTNFLEAKEGDLHRIEIPKFHMLH 381
Db 218 dmecikrralte---plvdtvgdnqlyvtdskllkmdiskmaagda-sfapfklyagr 272
OY 382 SGIVHGIAFWFDVAFTSISIMTWLSTAPTEPLTHWYORCLFQSPLEPAKAGDTLSGTCCL 441
Db 273 ndlhalvayfdvafstmcchkmgfstgpkrrathwkgvtvlyledvltlcegetltsmtl 332
OY 442 IANKROSYDISI 453
Db 333 agnkhnpdvdl 344
RESULT 12
ID G29548
G29548 standard; Protein; 306 AA.
AC G29548;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35176.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123568.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144329.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145228.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148568.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157173.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 14.0%; Score 443.5; DB 21; Length 306;
Best Local Similarity 36.4%; Pred. No. 6e-30; Indels 13; Gaps 6;
Matches 107; Conservative 51; Mismatches 123;

QY 163 MMDYRTGTGTOAIIQNTDFKRIIVLDVCGSGGILSFPAQAGARKIYAEASTWAOH 222
Db 1 mtkdvrtksygdvlyknhflkdkivldvgagtgllsfcaagahyavacsqmdt 60
QY 223 AEVLKSNMLTDRIIVIPCKVEEVSIP-EOVDIIIEPMGYMLFNERMLESYLHAK-KYL 280
Db 61 akeivksngfsdvltvlykqkieelvpkvdvliisemogyflllyenmltdvlyarnkwl 120
QY 281 KPSGNMPTIGDVHLAPFTDEQLYMEQFTKANRYPSPFHGVLSLRGAADVETFPQPV 340
Db 121 vdgglvlpckaslyvaled-----ahyrdkveifwddvygfdmsckrrate----pl 171
QY 341 VDTFD-IRITMAKSVKYYTNFLEAKEGDLHRIEIPFHMHSGLVHGLAFWFDVAFIGS 399
Db 172 vdtvgnqvltsklktkndiskmaagda-stfapiklvaqrndihahvaytvdstmc 230
QY 400 IMTWLSTAPTEPLHMYOVRCIFQSPFLAKAGDTLSGTCLLIANKRQSDISI 453
Db 231 hkkmgfstgpkrsathwkgvtvlyledvlticegetiltsmltqknknpdvdi 284
RESULT 13
ID W64204 standard; Protein; 360 AA.
XX W64204;
AC W64204;
DT 09-NOV-1998 (first entry)
XX Human HCP-1.
DE Human HCP-1.
XX HCP-1; IR1B4; Interferon receptor 1 binding protein 4; human.
OS Homo sapiens.
XX PN W09831796-A1.
XX PD 23-JUL-1998.
XX PF 15-JAN-1998; 98WO-US00671.
XX PR 15-JAN-1997; 97US-0035636.
XX PA (MCIN/) MCINNTS P A.
XX PA (YEDA) YEDA RES & DEV CO LTD.
XX PI Abramovitch C, Chebath JE, Revel M;
XX WPI, 1998-414096/35.

XX New isolated interferon receptor binding proteins - used to develop
 PT products for modulating sensitivity to interferon, e.g. in the
 PT treatment of tumours or for prolonging graft survival
 XX
 PS Example 5; Page 41-42; 64pp; English.

CC This polypeptide comprises human HCP-1, a protein originally
 CC identified on the basis of its ability to complement the Itel5
 CC mutation in yeast. A novel interferon receptor binding protein,
 CC IR1B4 (see W64202), of the invention shows homology to HCP-1.
 CC IR1B4 polypeptides and polynucleotides can be used to develop
 CC products for modulating sensitivity to interferon, e.g. in cancer
 CC therapy and for prolonging graft survival.

XX Sequence 360 AA;

Query Match 11.9%; Score 378; DB 19; Length 360;

Best Local Similarity 27.3%; Pred. No. 3.2e-24; Mismatches 102; Conservative 69; Mismatches 165; Indels 38; Gaps 8;

DB 2 enfatcl-angmslqppleevscg-----gaesekpnaedmtskdyfidsyah 49

OY 157 LSQOQNMADYVTRTGYVORAILQNHDTDFKDIIVDVGSGSILSFPAQAQARKIYAVEA 216

DB 50 fglheemlkdevrtlytnsmfhnrlfkdkvldvsgsgtqlcmfaakagarkvlgivc 109

OY 217 STMAQHAELVYKSNLNDRIYVIRGKYEVSLEPQVDIIIEPBGYMLFNRMLESYLHA 276

DB 110 ssisdvavkklvkaaklhnvrltkgkveeyelpye-kvaassasagvatastspostpc 168

OY 277 K-KYLKPSGMPPTIGDVHAPFTDEQLYMEQFTKANFRQPSFGHGVDSLALGAAYDE 334

DB 169 mpqtsvapgdlifpdrctlytated-----rqlkydklhwenvygdmsclkvaa-- 221

OY 335 YFROPVVDTRDIRILMAKSVKYTVNLEAKEGDLHRIE-----PEKFHMLHSGLVHG 387

DB 222 --keplavdvdpkqlv-----tnaclikevdiytvkvvedlftfscldgvkrndyha 272

OY 388 LAWEVDYAFIGSINTVLSNAPTEPLTHMYVRCLEFSGPLFAKAGDTLSGCLLIANKRO 447

DB 273 lvyvfnlefrchkrftgfspspepylhwkqvtlymedyltvktgeelfglgmprnakn 332

OY 448 SYDISIVAQVDQNG 461

DB 333 nrdldftdidfdkfg 346

RESULT 14

ID Y48597 standard; Protein; 297 AA.

XX Y48597;

XX 08-DEC-1999 (first entry)

DE Human breast tumour-associated protein 58.

XX Expressed sequence tag; EST; human; breast; cancer; gene therapy;

KW treatment; tumour; cytostatic; medicament.

XX Homo sapiens.

XX DE19813839-A1.

XX 23-SEP-1999.

XX 20-MAR-1998; 98DE-1013839.

XX 20-MAR-1998; 98DE-1013839.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PA Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 PI WPI: 1999-528981/45.
 XX N-PSDB: 233662.

XX Human nucleic acid sequences and protein products from tumor breast
 PT tissue, useful for breast cancer therapy -

PS Claim 25; 168; 188pp; German.

CC This invention describes novel human nucleic acid sequences from tumor
 CC breast tissue which have cytoskeletal activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC activity against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medicaments for gene
 CC therapy to treat breast cancer. Y48540-Y48617 represent protein fragments
 CC encoded by the expressed sequence tags described in the method of the
 CC invention.

XX Sequence 297 AA;

Query Match 9.5%; Score 299.5; DB 20; Length 297;

Best Local Similarity 29.4%; Pred. No. 1.3e-17; Mismatches 84; Conservative 51; Mismatches 122; Indels 29; Gaps 9;

OY 189 VLDVCGCGSILSF--FAQAQARKIYAVEASTMAQHAELV-KSNLNDRIY-VIRGKYE 244

DB 14 lmevscggaesekpnaedmtskdyfidsyahfglheemlkdevrtllidhvrltkgkve 73

OY 245 EVSLP-EGVDIIIEPBGYMLFNERMLESYLHA-KYLKPSGMPPTIGDVHAPFTDEQ 302

DB 74 eelprvekvdlileemwylcylfiesmnlvlyarkvlapdglifpdrctlytated-- 131

OY 303 LYMEQFTKANFRQPSFGHGVDSLALGAAYDEYFROPVVDTRDIRILMAKSVKYTVNLE 362

DB 132 --rqlkydklhwenvygdmsclkvaa--keplavdvdpkqlv-----tnacl 177

OY 363 AKEGDLHRIE-----PEKFHMLHSGLVHGSLAWEPDVAFIGSINTVLSNAPTEPLTH 415

DB 178 lkeyvdiytvkvvedlftfscldgvkrndyvalvynlefrchkrftgfspspepyh 237

OY 416 WYVRCLEFSGPLFAKAGDTLSGCLLIANKROSYSISIVAQVDQNG 461

DB 238 wkqvtlymedyltvktgeelfglgmprnaknrdldftdidfdkfg 283

RESULT 15

ID Y26923 standard; protein; 122 AA.

XX Y26923;

XX 07-DEC-1999 (first entry)

DE Amino acids 58-179 of protein hmtl_yeast.

XX Classification; biological function; hierarchical organization;

KW database; cluster; nucleic acid; glycoprotein; sugar; immunoglobulin;

KW chemical; measures of similarity; Smith-Waterman; FASTA; BLAST;

KW G-protein; Ras; biosynthesis.

XX Saccharomyces cerevisiae.

XX WO939174-A2.

XX 05-AUG-1999.

XX 29-JAN-1999; 99WO-IL00057.

XX 29-JAN-1998; 9805-0072977

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM
XY

PI Linial M, Linial N, Tishby N, Yona G;
YY

DR WPI; 1999-458952/38.

PT Automatic method for classifying molecules having similar biological
PT function, - used for identifying similarity among, e.g. proteins,
PT nucleic acids, glycoproteins, complex sugars or immunoglobulins
XX

PS Disclosure; Flg 8; 78pp; English

CC The invention relates to an automatic method (AM) for classifying
CC molecules having similar biological function by: (a) creating a
CC hierarchical organization of the molecules in a database, where groups
CC of clusters are identified using local consideration resulting in
CC related clusters, and (b) determining the position of a selected
CC molecule based on the hierarchical organization of (a), where selected
CC molecules of similar biological function are classified. The method can
CC be used for classifying molecules such as nucleic acids, DNA or RNA,
CC polypeptides, proteins, peptides, glycoproteins, complex sugars,
CC immunoglobulins or chemicals. Since the method may comprise applying
CC standard measures of similarity e.g. Smith Waterman, FASTA or BLAST,
CC these can be used to provide hierarchical organization within known
CC families, e.g. the small G-protein/Ras super-family.
CC This protein forms a part of a protein from a hierarchy of methylase
CC or methyltransferase proteins and is designated mnt1_yeast.
CC The sequence is used to illustrate the invention.

SQ Sequence 122 AA;

Query Match	8.48;	Score 267;	DB 20;	Length 122;
Best Local Similarity	48.38;	Prod No. 3	30-15;	

Matches 58; Conservative 24; Mismatches 36; Indels 2; Gaps . 2;

Db 1 kdkivldvgcgtglismfaakbhgakhvlgcmssliemakelveingfsdsktllrgkle 60

QY 245 EVSLP-EQVNIISSEPMGYLMENRMLESYLHAK-KYLKPSGNNPPTIGVHLAPPTDEQ 3022

Db 61 dvhlfpkvdlllsewmyfllyesmndtvlyardhyivegllifpdkcsihlagledsq 1200

Db 61 dvhlfpkvdllsewmgyllyesmdtvl yardhy lveggllfpdkcsihlagledsq 120

Search completed: June 20, 2001, 09:57:38
Job time: 152 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 09:55:06 ; Search time 22.66 Seconds
(without alignments)
2043.871 Million cell updates/sec

Title: US-09-464-377-2

Sequence: 1 MAANAATVAGPGAGSAGVAG.....PAISMASPSIPTNTHMIGS 608

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	498.5	15.7	390	2 T09914	protein-arginine N
2	496.5	15.7	376	2 T52248	protein-arginine N
3	486	15.3	348	2 S45890	ODP1 protein - yea
4	477.5	15.1	383	2 A86182	hypothetical prote
5	472.5	14.9	339	2 T50263	probable arginine
6	467.5	14.8	366	2 F84579	hypothetical prote
7	449.5	14.2	428	2 T26447	hypothetical n-methyl
8	374	11.8	472	2 T40755	related to protein
9	349	11.0	521	2 T49355	hypothetical prote
10	172	5.4	630	2 T26236	flission yeast Skbl
11	170.5	5.4	637	2 T03842	hypothetical prote
12	163	5.1	670	2 T10666	conserved hypotet
13	146	4.6	259	2 C64481	hypothetical prote
14	138	4.4	680	2 T25146	hypothetical prote
15	123.5	3.9	271	2 T41026	probable methyltra
16	122.5	3.9	638	2 T53169	cytochrome 2 - hu
17	120.5	3.8	208	2 F75025	methyltransferase
18	120.5	3.8	1077	2 S45395	hypothetical prote
19	119.5	3.7	245	2 S46002	probable membrane
20	118	3.7	245	2 D82221	3-demethylubiquino
21	115	3.6	296	1 T64105	ribosomal protein
22	114	3.6	387	2 A81324	cyclopropane-fatty
23	114	3.6	819	2 T08745	probable RNA helic
24	112.5	3.6	207	2 D71210	hypothetical prote
25	112.5	3.6	267	2 A72338	hemolysin - Thermo
26	112.5	3.6	529	2 T23431	hypothetical prote
27	112.5	3.6	1997	2 T30874	virginiamycin S sy
28	111	3.5	258	2 H70367	hemolysin - Aquife
29	111	3.5	1032	2 T34433	hypothetical prote

30	111	3.5	1755	2 S69951	TyB protein - yea
31	111	3.5	2232	2 T34434	hypothetical prote
32	110	3.5	197	2 H64415	hypothetical prote
33	109.5	3.5	311	1 C69952	probable ribosomal
34	109	3.4	1755	2 S69866	TyB protein - yea
35	108	3.4	356	2 A47128	carbamimycin 4-O-m
36	107.5	3.4	295	2 D82341	ribosomal protein
37	107.5	3.4	1289	2 T18212	parasporeal crystal
38	107	3.4	144	2 T33109	hypothetical prote
39	106.5	3.4	690	2 JC4185	protease II (EC
40	106.5	3.4	245	2 B86738	hypothetical prote
41	106.5	3.4	967	2 S66852	hypothetical prote
42	106	3.3	216	2 A72345	conserved hypotet
43	106	3.3	314	2 T43740	probable ribosomal
44	105.5	3.3	771	2 H84845	hypothetical prote
45	105	3.3	251	2 B64048	hypothetical prote

ALIGNMENTS

RESULT 1
T09914
protein-arginine N-methyltransferase (EC 2.1.1.23) - Arabidopsis thaliana
N:Alternate names: protein T16L4.20
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 05-May-2000
C:Accession: T09914
R:Bayan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16897
A:Accession: T09914
A:Molecule type: DNA
A:Residues: 1-390 <BEV>
A:Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.20
A:Experimental source: cultivar Columbia; BAC clone T16L4
C:Genetics:
A:Gene: paml; ATSP:T16L4.20
A:Map position: 4
A:introns: 80/1; 83/3; 135/3; 157/1; 204/3; 234/1; 272/3; 323/1
C:Keywords: methyltransferase; S-adenosylmethionine

Query Match	15.7%	Score 498.5	DB 2	Length 390
Best Local Similarity	36.4%	Pred. No. 1.3e-29		
Matches 125	Conservative 55	Mismatches 142	Indels 21	Gaps 8
QY	114	TPHDFCSYNNIKTRGHTLERSVSERTESSAVOYFOFYGLSOOQNMADYVATGY	173	
DB	44	TPODESMF-----DAGESADTAETVDTDT--TSADYFDSYSHFGIHEMLDVAETKTY	95	
QY	174	ORALIQNTDFDKITLVLDVCGSGGILSFFAQAQARKIYAVASTMAQAEVLVKSNNLT	233	
DB	96	QNVITQNFELIKDKITLVLDVGAAGTILSLFCARAGAAHYAVSCQADAKETVRANGS	155	
QY	234	DRIVVIPGKVEEVSIP-EQVDIIISEPMGYMLFNEMLESYLAH-KYLPKSGNMFPITG	291	
DB	156	DVITVLKKGKIEIELEPTPKVDVITISEMMGYFLFEMLDLSVLYARKKMLEGGVLPDKA	215	
QY	232	DVHLAPFTDEQLYMQFTRKANFRYQPSFGVLSALRGAAVDYFRQPVVOTFDIRILMA	351	
DB	216	SLHLTAID-----SEYKEDKIEFNWSYVGFDMSCIKKAKME---PLVDVDOQNIAT	266	
QY	352	KS-VKVTNPLFAKGGDHRIRIEPKRFHMLHGLVGLAFWDAFVIGSIMVNSTAPT	410	
DB	267	DSRLKTKMDISKSSGDA-SFAPKVLAQRNDYHALVAIYDVASFTHCKLLGSGTGK	325	
QY	411	EPLTHMYVRCFLFQSPLEPAKAGDTLSGTCCLTANKROSDYIS	453	
DB	336	SRATHWKQTVLYLEVDVLTICBETITGTMSVSPNKKNPDDIDI	368	
RESULT	2			

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86441; MUID:21016719
A:Accession: A86182
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1583 <STO>
A:Cross-references: GB:AE005172; MID:g7211979; PIDN:AAF04050.1; GSPDB:GN001411
C:Genetics:
I:Map position: 1

Query Match	15.1%;	Score 477.5;	DB 2;	Length 383;
Best Local Similarity	33.6%;	Pred. No. 4.7e-28;		
Matches 125; Conservative	66;	Mismatches 124;	Indels 57;	Gaps 12

```
QY      143 EESSAVQVPEFVEGYLSQQNNMADYVRIGTYQRALILNHTDFKDKIVLVDGCGSGILISFF 202
        :| || | :| :||| | | | | | :| | :| | | | | | | | :
Db       28 KEVDYAQVFCTYSFLYHKDMDSRVRDAYFNANVFQNKHHEGSTVLVDVGSGILAIIW 87
```

Qy 203 AAAGARKIIVAEASTMAQHAEVLVKSNNLTDRIVIPCKVEEYSLPEVDIIISEPMGY 262
:|||||:||||: || |::||| : || |::|||:|||| ||
Db 88 SAQGARKVVAVEATKADHARALVKANNLDHIYEVIIEGSEVEDISLPEKVDIISEMMGY 147

```

Oy 263 MLENERMLESTYLHAK-KYLKPSGNNPEPTIGDVHAP-----FT 299
    | | | : | : : ||| | | | : | | |
Db 148 FLARESMFDSYISARDRLKPGVMYPSHARMLAPIKSNADRRKNDPFGAMADWNFS 207

```

```
Oy      300 DEOLMEQFTKANFRYQPSEHGVDLSALRGAAYDE--YFQRPVV-DTFDIRIMAKSV   354  
         ||          . : ::|||: |           :| :|: |       :|:  
Db     208 DE-----IKSYIGVDMGVTITKPFPEEQEKIYTQTAMNNDLNQQIIIGPPTI    253
```

```
Oy      355 KYTVNFELEAKEGDLH--RIEIPKPFHMLHSGLVHGIAFWEDVAFIG-----SMTWVLST 407
          ::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
Db      254 VKENDCLTASYSEIEFVRNSNTSVINNETHRLC-GFGGTFDVOEGSRKKEDPAQOEIELTT 312
```

QY 408 APTPE-LTHWYQVRCLLEQSPLEAKAGDTLSTGCLLIANKRQSY---DISIYAQVDQTGSK 4633

Db 313 APSECHCTHKGQOVFTMSNPINVEEGDILN-LGILMSRKENHRLMEIELNCEIKASGN 3711

QY	464	SSNLDDLKNPF	475
		:	
Db	372	PKE--SEKKTYP	381

RESULT 5
150263

C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C:Accession: T50263

submitted to the EMBL Data Library, December 1999
A:Reference number: Z25052
A:Accession: T50263

A: Molecule type: DNA
 A: Cross-references: EMBL:AL133498; PTDN:CAB63498.1; GSPDB:GN00066; SPDB:SPAC890.07c
 A: Molecule type: DNA
 A: Cross-references: EMBL:AL133498; PTDN:CAB63498.1; GSPDB:GN00066; SPDB:SPAC890.07c

A:Map position: 1
 A:Gene: SPDB:SPAC890.07c
 C:Genetics:
 A:Experimental source: Strain 9/24(-); COSmid c890

Query Match 14 98: Score 473 5: DB 3:
A: INCONS: 16/3; 18/1; 22/2; 60/3; 82/3; 181/1; 292/3

150. VPOEWCVI GCGGMMQDYRMETMVKQKRIYQNMDEKVTTH SVQCGCCGCTT CHTTATQCAID 300
Best Local Similarity -36.9%; Pred. No. 9.3e-28;
Matches 114; Conservative 49; Mismatches 131; Indels 15; Gaps

Db 19 YFDSYSHWGTHHEMLKNDVRLTSLYDAIMQNPHELRDKIYLDVGCGTGLSNFCARAGAK 78

[illegible]

RESULT 6
F84579
probable arginine N-methyltransferase [Imported] - Arabidopsis thaliana

C:\Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:\Accession: F84579
Rülin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.;

Euss, D.; Niehan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999

Article: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Accession: F84579
A:Status: Preliminary
A:Molecule type: DNA

A:Cross-references: GB:AE002093; NID:g3687250; PIDN:AAC62148.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g19670

Query Match

05	best_local_similarity	55.98%	file:	no.2.3c.27;						
	Matches	112;	Conservative	54;	Mismatches	133;	Indels	13;	Gaps	6;
145	SSAVOYFOFYGLSLOOONMODIVRTGTYORAILONHTDFDKRILVDPGCGSILSFFAA	204								

0v 205 OAGARKIYAEASTMAOHAEEVLKSNNTDRIYVIPGKVEEVSIP-FOVDIIISFPMGYM 26

Db 43 TSADYYPFDSISHFGIHEEMLKDVVRITSQDVLYKKNFLIKDKIYLDVGAGIGILSLFCA 10

```

0v      26d  IJNFRMIESYIHAK-KYIKPSCNMFPPTGDUVHIAETDNEOTUMFOETKANEBRYOPSEHCY 3222
Db      103 KAGAAHVIAVECSQADTAKEIVKSNGESDVITVLKGIEIELPYPKVDVIISEMMGYF 165d

```

```

163 LLYENMLDFVLYARNKWLVDGIVLPKASLYTAIED-----AHYKDKVEFMDVYGF 217

```

```

218 DMSICKRRAITE-----PLVDYWDGNGLYTDSKLKLTMDSKMAAGDA-SFTAPFKEVAQR 272Z

```

273 NDHIALVAFDVSTMCCKKNGFSTGPKSRATHWQIVLYLEDVLTICESTITGSMFI 332

Db 333 A Q N K K N P R D V D I 344

RESULT 7

T26447
 hypothetical protein Y113G7B.17 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T26447
 R:Jennard, N.
 submitted to the EMBL Data Library, September 1999
 A:Reference number: Z20215
 A:Accession: T26447
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1348 <NID>
 A:Cross-references: EMBL:AL10477, NID:e1542121; PIDN:CAB54335.1; CESP:Y113G7B.17
 A:Experimental source: clone Y113G7B
 C:Genetics:
 A:Gene: CESP:Y113G7B.17
 A:Introns: 90/3; 297/3

Query Match 14.2%; Score 449.5; DB 2; Length 348;
 Best Local Similarity 32.9%; Pred. No. 5,1e-26;
 Matches 115; Conservative 62; Mismatches 144; Indels 29; Gaps 8;

QY 125 LKTCRGITLRSVSEETEE-SSAVOYFOFYGYLSQOONMADYVRKGTGRALIONHTD 183
 DB 1 MSTENGSADAPVAPAPAKELTSKDYFDSTAHGHEMLKDEKRTTYRNSIYHNSHL 60
 QY 184 FKDKIVADGCGSGILSFFAQAQARKIYAVEASTMAQHAELVKSNNLTDRIYVYIPKV 243
 DB 61 FKDKVAVDVSGSITSLMFAKAKAKKVFAMFPMALTSKTIADNNLDHIVEYIAKV 120
 QY 244 EKV-SLP---EQVDIIISEPGYMLFNERMLESTYIAK-KYIKPSGNPFPTIGDVHIAFP 298
 DB 121 EDVHELEGGIEKVDIIISEMNGYCLFYESMLNTVLAADRMILAPNGMLFPDKARLYVCAI 180
 QY 299 TDEDLVMEQTKANFRQPSFHGVDLSALRGAAVDEYRQGVNPTFDRIIMAKSVKTYV 358
 DB 181 ED-----RQYKEDIHWDSDYGFNMSAIKNVAI---KEPLVDIVD-----NAQVNT 224
 QY 359 NPLEAKEGDLHRIE-----PREFHMLHSLVGLAFMFDAVFISIMTWLSTAPTE 411
 DB 225 NNCILKAVDLYTVKIEDLTFKSDPKLNCTRSDYIQAFYFTFVESKCHKKTGFSTGPDV 284
 QY 412 PLTHVYVRCLEFOSPLFAKAGDTLSGTCCLIANKRQSTYDISIVAQVDQTG 461
 DB 285 QYFHWKQTVFLKDALTVKKGSEITGSEMAPNKNNERDLINISFDRKG 334

RESULT 8
 T40755
 arginine n-methyltransferase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T40755
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
 submitted to the EMBL Data Library, March 1998
 A:Reference number: Z21948
 A:Accession: T40755
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-472 <LYN>
 A:Cross-references: EMBL:AL020772; PIDN:CA117825.1; GSPDB:GN00067; SPDB:SPBC8D2.10C
 A:Experimental source: strain 972h-; cosmid c8D2
 C:Genetics:
 A:Gene: SPDB:SPBC8D2.10C
 A:Map position: 2

Query Match 11.8%; Score 374; DB 2; Length 472;
 Best Local Similarity 28.0%; Pred. No. 3,6e-20;
 Matches 133; Conservative 65; Mismatches 169; Indels 108; Gaps 16;

QY 30 VSVFP-GARLLITGDANGEIQRHAE-QQALRLLEVAPGADAGIALYSHEDVCYFKCSVSR 87

DB 61 ISVLPDSDVLSLGD---ELDSFEDDNTLELEVENPAD-----VSK 99
 QY 88 ETEGSRVROSEFIITLGCNSVLIOGATPHDCFSNIIKTCRGHTLERSVFSERTSESSA 147
 DB 100 DAELIKKLQ-----NOLI-----SOLEIKDKM-NELTSQTDQLSV 138
 QY 148 V-----QYFOFYGYLSQOONMADYVRKGTGRALIONHTDEKDIVLDVGGSGIL 199
 DB 139 TPKRADNSYFEESYAGNDIHFLMLNDSVRTEGDFYVHNKHIFAKGTVLDDVCGGTGL 198
 QY 200 SFFAQAQARKIYAVEASTMAQHAELVKSNNLTDRIYVYIPKVEVSLP-EQVDIIISE 258
 DB 199 SMFCQAKAKKRYAANDSDIIOAMISNAFENGLAQIIFIRKIDISLPGVKVDIIISE 258
 QY 259 PMGYVLFNERMLESTYIAK-KYIKPSGNPFPTIGDVHIAFPDEOLYMEQF----- 309
 DB 259 WKGVALTFESMIDSVLVARDRFLAPSGIMAPSETVLATNTLELPEIDFMSDVYGF 318
 QY 310 KANFRQPSFHGVDLSALRGAAVDEYRQPVV-DTFDIRIIMAKSVKTYVNFLEAKGDL 368
 DB 319 KNGMKDASYGVSVOVPPQTYVN---AKPVYAFRNHRTCKVQDVSEF----- 364
 QY 369 HRIEPEFHLHSLVGLAFMFDAVFISIMTWLSTAPTEPL----- 413
 DB 365 ---SFLSIIDNEGRLAFLTFMPTVF-----TKRQPIPEAIDEACGFTTGPO 411
 QY 414 ---FWYVRCLEFOSPLFAKAGDTLSGTCCLIANKRQSTYDISIVAQVDQTSKSS 465
 DB 412 GTPHMKQCVLLRRNRPLOKGTREYGTISPSKNNKNNRDDISVHWNVNGADS 466

RESULT 9
 T49355
 related to protein arginine N-methyltransferase 3 [imported] - Neurospora crassa
 N:Alternate names: protein B1D1.90
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 R:Schulte, U.; Hohelsel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatu
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49355
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-521 <SCH>
 A:Cross-references: EMBL:AL355927; GSPDB:GN00116; NCSP:B1D1.90
 A:Experimental source: BAC clone B1D1; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B1D1.90
 A:Map position: 6
 A:Introns: 33/1

Query Match 11.0%; Score 349; DB 2; Length 521;
 Best Local Similarity 29.9%; Pred. No. 3,1e-18;
 Matches 95; Conservative 53; Mismatches 106; Indels 64; Gaps 10;

QY 143 EESSAVOYFOFYGYLSQOONMADYVRKGTGRALIONHTDEKDIVLDVGGSGILSFF 202
 DB 168 KEGADYTFESTAHNDIHETMLKDYVRKEAYVDFIYONKDLFAGKVVLDIGCTGLISNF 227
 QY 203 AAQAQARKIYAVEASTMAQHAELVKSNNLTDRIYVYIPKVEVSLP-EQVDIIISEPMG 261
 DB 228 CAKAKAQOVIADVRSSEIIDKARENITANGLSVITLKRIREVILIPVEKVDIYSEMMG 287
 QY 262 YMLFNERMLESTYIAK-KYIKPSGNPFPTIGDVHIAFPDEOLYMEQFKANFRYQPSH 320
 DB 288 YCLLEAMINSLVLMARDYLAPOGLVPSHGNMWIAP-VSEQDYIAEYVD---FWRDYV 342
 QY 321 GVDLSALRGAAVDEYRQPV-----VDFDIRIIMAKSVKTYVNFLEA---KEG 366
 DB 343 GFDMAVMO-KGIYEDCRMEYRPAETVCCGPASFGLLDHFVAVEDLVFTAKQASAPDKA 401

OY 367 DLHRIEIPKFMHLSGLVHGLAFWEDVAF-----IGSIMTVML----- 405
 DB 402 ESR-----DGLVWMDVFFANRNVDSEIKIDTKAOWVAETAGKGDKD 445
 OY 406 -----STAPTEPLTHWYQ 418
 DB 446 ARVAFITGPPEPTHMRO 463
 RESULT 10
 T26236
 hypothetical protein W06D4.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T26236
 R:Kershaw, J.
 submitted to the EMBL Data Library, November 1998
 A:Reference number: 220178
 A:Accession: T26236
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-630 <full>
 A:Cross-references: EMBL:AL034364; NID:e1356156; PIDN:CAA22252.1; GSPDB:GN00019; CESP:W06D4
 A:Experimental source: clone W06D4
 C:Genetics:
 A:Gene: CESP:W06D4.4
 A:Map position: 1
 A:Introns: 30/2; 72/1; 200/1; 296/3; 359/3; 465/3; 553/2
 Query Match 5.4%; Score 172; DB 2; Length 630;
 Best Local Similarity 23.8%; Pred. No. 7.3e-05;
 Matches 63; Conservative 45; Mismatches 99; Indels 58; Gaps 10;
 OY 162 NMADYRTGYORAIL-----QNHTEPKDKI-VLDVCGSGSITSFPAAGARRIYAVE 215
 DB 35 DMILDFPRNDKFLAGLTKTIAEKHEKNDGKGVHVDIGTGGLSLMARECADVTALE 94
 OY 216 A-STMAOHAELVYKSNLTDRIIVIPGKVEEYSL--PEOVDIIEPAGYMLFNERMES 272
 DB 95 VFKMGDCARRITTSNPSDSKITIYISERSTYSQIGSRADITVAEVDTELIGGALRT 154
 OY 273 YLHA-KRYLKSGNMFFITGDVHLAFTDEOLYMEQFTKANFRYOPSPHG----VDLSAL 327
 DB 155 FKELERLAKGCRVVPSTGVNVIYVPSHLTKM-----FNDIPRLNGEDEDEPLGRG 207
 OY 328 RGAANDVEFROPVYDFTDIRILMAKSVYTVNFEAKEGDLHRIIPKFMHLR----- 381
 DB 208 SGTAA-----VDDVQISEKKTHERF-----ELSEPIVAKRFDFEHEEKIIF 248
 OY 382 -----SGLVHGLAFWFDV 394
 DB 249 DESFVREAVAHSSGTIDALLMMWDI 273
 RESULT 11
 T03842
 fission yeast Skh1 protein homolog - human
 C:Species: Homo sapiens (man)
 C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 16-Feb-2001
 C:Accession: T03842
 R:Marcus, S.
 submitted to the EMBL Data Library, July 1997
 A:Description: SKB1Hs, a human homolog of the fission yeast skh1 gene.
 A:Reference number: Z15114
 A:Accession: T03842
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-637 <NAR>
 A:Cross-references: EMBL:AF015913; NID:g2323409; PIDN:AA06581.1; PID:g2323410
 C:Genetics:
 A:Gene: SKB1Hs

C:Superfamily: Schizosaccharomyces pombe negative regulator of mitosis skh1
 Query Match 5.4%; Score 170.5; DB 2; Length 637;
 Best Local Similarity 22.8%; Pred. No. 9.6e-05;
 Matches 94; Conservative 58; Mismatches 152; Indels 109; Gaps 19;
 OY 99 FITTLCNSVLIQFATPHDFCSFYNIKTCRGHTLERSVSEPTTESSAVQYQFYGLS 158
 DB 264 FITT-GTN-----HSEKEFCSYIOLYLE-----YLSQRPPPNAYELFA-KGYED 306
 OY 159 QOONMMQ-----DYVRTGYORAIL-----QNHTEPKDKIIVDVCG 194
 DB 307 YLQSPLOPLMDNESQTYVEFKDPITKISQYQATIKCLLDRYPEEKDNNVLAWLGA 366
 OY 195 GSGIL---SFFAAOAGAR-KIYAVEASTMAQAEVLVKSNN---NLTDRIVVIPGKVE 245
 DB 367 GGGPLVNASLRKAKQADRITKILAYEKRNPA---VYTLNMQFEEMGSOYIVYSSDMRE 422
 OY 246 VSLPROVDIIESEPQYMLFNERMLESYLHAKKYLKPSGNMFFITGDVHLAFTDEOLYM 305
 DB 423 WVAPEKADITIVSELLGSFADNELSPECLDGAQHFLDDGVISIGETSTFLAPLSSSKLVN 482
 OY 306 E-----OPTKANFRYQPSFGVDSALRGAAVDEY-FROPVYDFTDIRILMA 351
 DB 483 EVRACREKDRDEAOFEMEVYVRLHNFH--QLSAPOPCFTFSHPNRDPMID----- 531
 OY 352 KSVKTYVNLKAEKGLHRIEIPKFMHLSGLVHGLAFWEDVAFGSIYTVLSTAPTE 411
 DB 532 -NNRYCT-----LEPPEVNT-----VLHGFAYFETVLYQDI-----TLSIR 568
 OY 412 PLTH-----WYQVRLCFOSPLFAKADTLTSGTCLLIAN-KROSYDISIYAOV 457
 DB 569 PETHSPGMSWPIPLPIKQIPITVRGQITICVAFMCSKSKKWTYEMAYAPV 621
 RESULT 12
 T10666
 hypothetical protein F6E21.40 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Feb-2001
 C:Accession: T10666
 R:Bayan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Ban
 submitted to the protein sequence database, June 1999
 A:Reference number: Z16533
 A:Accession: T10666
 A:Molecule type: DNA
 A:Residues: 1-670 <BEV>
 A:Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.40
 A:Experimental source: cultivar Columbia; BAC clone F6E21
 C:Genetics:
 A:Gene: ATSP:F6E21.40
 A:Map position: 4
 A:Introns: 47/3; 87/1; 123/3; 203/3; 230/3; 255/3; 284/3; 305/1; 335/3; 347/3; 370/3;
 C:Superfamily: Schizosaccharomyces pombe negative regulator of mitosis skh1
 Query Match 5.1%; Score 163; DB 2; Length 670;
 Best Local Similarity 22.9%; Pred. No. 0.00038;
 Matches 80; Conservative 53; Mismatches 145; Indels 72; Gaps 11;
 OY 131 HTLERSV--FSRTESSAVQYQFYGYSOQONMM----- 164
 DB 293 HNLQKPLDSSSGTCKENPLRITLDVYAYILFQKMEISSEDERIELGRTDLOAPLOPLMDN 352
 OY 165 -----QDYVRTGYORAILQNHTE-FKDK-----IYLDVCGSGGIL---SFFA 203
 DB 353 LEAQTYETFERDSVXYIQORAVAKALVDYRDEKASELTLYLMVYAGARGPLVRSALQ 412
 OY 204 AQAGAR--KIYAVEASTMA-QHAELVYKSNLTDRIIVIPGKVEEYSLPEQVDIITSEPM 260
 DB 413 AETDRKLKLYAAVEKNPNNAVVTLLHNLVKMEGMEVDVTTISCDMRFWNADEQADITIVSELL 472

Thu Jun 21 08:07:40 2001

us-09-464-377-2.rpr

Page 7

Db 186 LLA 188

Search completed: June 20, 2001, 09:58:04
Job time: 178 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 09:57:16 ; Search time 13.03 Seconds
(without alignments)
1598.414 Million cell updates/sec

Title: US-09-464-377-2

Sequence: 1 MAAAAATAGVPGAGSAGVAG.....PAISMASPMISPTNTHGYS 608

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 segs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494.5	15.6	353	1 ANM1_RAT	Q63009 rattus norv
2	494.5	15.6	371	1 ANM1_MOUSE	Q9J110 mus musculu
3	486	15.3	348	1 HMT1_YEAST	P38074 saccharomyc
4	484.5	15.3	334	1 ANM4_HUMAN	Q9NR22 homo sapien
5	481.5	15.2	361	1 ANM1_HUMAN	Q99873 homo sapien
6	475	15.0	448	1 ANM2_MOUSE	Q99873 homo sapien
7	474.5	15.0	512	1 ANM3_HUMAN	Q60678 mus musculu
8	472.5	14.9	339	1 ANM1_SCHPO	Q60678 homo sapien
9	472	14.9	528	1 ANM3_SCHPO	Q60678 homo sapien
10	461	14.6	433	1 ANM2_HUMAN	Q70467 rattus norv
11	123.5	3.9	271	1 COO3_SCHPO	P55345 homo sapien
12	122.5	3.9	638	1 K220_HUMAN	Q01546 s hexapreny
13	120.5	3.8	1077	1 YBKL_YEAST	P38167 saccharomyc
14	119.5	3.8	827	1 YBY3_YEAST	P38274 saccharomyc
15	115	3.6	296	1 PRMA_HABIN	P44402 haemophilus
16	110	3.5	197	1 Y928_MERUA	Q58338 methanococc
17	109.5	3.5	311	1 PRMA_BACSU	P54460 bacillus su
18	108	3.4	355	1 CMAT_STRPE	Q06528 streptomyce
19	107.5	3.4	1289	1 C5AB_BACUD	Q45753 bacillus th
20	107	3.4	690	1 PTRB_MORLA	Q59536 moraxella t
21	105	3.3	251	1 Y095_YARLI	Q57060 haemophilus
22	104.5	3.3	509	1 HOYL_YARLI	Q99160 yarrowia li
23	103.5	3.3	300	1 PRMA_SYNY3	P73820 synechocyst
24	103.5	3.3	705	1 SYNL_HUMAN	P17600 homo sapien
25	103	3.3	656	1 DNLJ_HELPY	Q25338 heliobacte
26	103	3.3	1281	1 YLBS_CAEEL	Q25338 heliobacte
27	103	3.3	1755	1 YJZ9_YEAST	P47100 saccharomyc
28	102.5	3.2	191	1 YK51_ARCFU	Q28228 archaebact
29	102.5	3.2	389	1 SERI_BOWMO	P07856 bombyx mori
30	102.5	3.2	656	1 DNLJ_HELPY	Q25338 heliobacte
31	102	3.2	880	1 BRCA_DROME	Q24206 dirosophila
32	100	3.2	1472	1 A2MG_RAT	Q62328 rattus norv
33	99.5	3.1	298	1 PRMA_CLOAB	P45558 clostridium

34	99.5	3.1	312	1 PRMA_STAUV	P45557 staphylococ
35	99	3.1	552	1 DCIP_ENTCL	P23234 enterobacte
36	98.5	3.1	596	1 HMEN_ANOGA	Q02491 anopheles g
37	98.5	3.1	3712	1 ACYS_CEPAC	P25464 cephalospor
38	98	3.1	477	1 ANGT_MOUSE	P11859 mus musculu
39	98	3.1	1505	1 CDE_HUMAN	P39880 homo sapien
40	98	3.1	1755	1 YJZ7_YEAST	P47098 saccharomyc
41	97.5	3.1	985	1 AGLU_ASPOB	Q12558 aspergillus
42	97.5	3.1	1148	1 ICEK_PSEEX	Q30611 pseudomonas
43	97.5	3.1	706	1 PLB2_YEAST	Q03674 saccharomyc
44	97	3.1	750	1 MEI2_SCHPO	P08965 schizosacch
45	97	3.1	1286	1 AIDA_ECOLI	Q03155 escherichia

ALIGNMENTS

RESULT	ID	ANM1_RAT	STANDARD:	PRT:	353 AA.
AC	063009:				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	PROTEIN ARGININE N-METHYLTRANSFERASE 1 (EC 2.1.1.-)				
GN	HMT112 OR PRMT1.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96278999; PubMed=8663146;				
RA	Lan W.-J., Gary J.D., Yang M.C., Clarke S., Herschman H.R.;				
RT	"The mammalian immediate-early TIS21 protein and the				
RT	leukemia-associated Btg1 protein interact with a protein-arginine				
RT	N-methyltransferase."				
RL	J. Biol. Chem. 271:15034-15044(1996).				
CC	-1- FUNCTION: METHYLATES (MONO AND ASYMMETRIC DIMETHYLATION) THE				
CC	ARGININE RICH DOMAIN (CAN METHYLATE HISTONAL AND HISTONES).				
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).				
CC	-1- TISSUE SPECIFICITY: UBIQUITOUS.				
CC	-1- SIMILARITY: BELONGS TO THE PROTEIN ARGININE N-METHYLTRANSFERASE				
CC	FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/				
CC	or send an email to license@sdb.ch).				
CC	-----				
DR	EMBL; U60882; AAC52622.1; -				
KW	Transferase; Methyltransferase; Nuclear protein.				
SQ	SEQUENCE 353 AA: 40522 MW; EDBB0587784CS272E CRC64;				
Query Match	15.6%; Score 494.5; DB 1; Length 353;				
Best Local Similarity	34.0%; Pred. No. 2.3e+29;				
Matches 115: Conservative 64; Mismatches 128; Indels 31; Gaps 7;					
OY	139 SERTESSAVO-----YQFYGYLSQOQNMADYRTTYQRAIQNTDFDKIVLDV 192				
DB	1 1				
OY	18 AESSKPAVMEWTSKDYFDSYAHFGIHEMLKDEKRTLYRNSMHNHLFDKVLVDV 77				
DB	1 1				
OY	193 GCGSIIILFFPAOAGARIVYAEASTMAOHAEVLVYNNLTGRIYVYIPGKVEEVSIP-DO 251				
DB	1 1				
OY	78 GSGTILILFFPAKAKARVYIGECSSISDYAVKIKANKLDHVVYITIKKVEEVEELPVKR 137				
DB	1 1				
OY	252 VDIISPMGYMLFERNMLESYIAHK-KYLRKSGNMFPITIGDVHLAPFTDEOLYMEQFTR 310				
DB	1 1				

```

Db 138 VDIISSEMGYCLFYESMNLTVLHARDKMLAPDGLIFPRATITVTAIED-----RQYKD 192
OY 311 ANFRIPSPSHGYDLSALRKAAVDEFRQPVDTFDIRILMAKSVKTYVNFLEAKEGDLHR 370
Db 193 YKIHMMENYVYGDMSCKIKVAI-----KEPLVVDVDPKQVLV-----TNACLIKEVDIYR 241
OY 371 IETI-----PKRFHMLHSGLVHGLAFWFDVAFISIMVWLTAPDEPLTHMYQVRCLE 423
Db 242 VVEDLTFTSPFCLOVKKNDVYHALVAFNIEFTRCHKRTGFTSPSPSYTHMKQVTFYV 301
OY 424 OSPLFAKAGDTLSCGCLLIANKKROSYDISIVAOVDQTG 461
Db 302 EDYLVTKGEBIEFTIGMRPNKNNDDLTFTDLDFKG 339

RESULT 2
ANML_MOUSE STANDARD: PRT: 371 AA.
AC 09JTF0;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTEIN ARGININE N-METHYLTRANSFERASE 1 (EC 2.1.1.-).
GN HRMTL12 OR PRMT1 OR MRMT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20307889; PubMed=10848611;
RA Pawlax M.R., Scherer C.A., Chen J., Koshon M.J., Ruley H.E.;
RT "Arginine N-methyltransferase 1 is required for early postimplantation
mouse development, but cells deficient in the enzyme are viable.";
RL Mol. Cell. Biol. 20:4859-4869(2000).
CC -1- FUNCTION: METHYLATES (MONO AND ASYMMETRIC DIMETHYLATION) THE
CC GUANIDINO NITROGENS OF ARGINYL RESIDUES PRESENT IN A GLYCINE AND
CC ARGININE-RICH DOMAIN (CAN METHYLATE HNRNP1 AND HISTONES).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE PROTEIN ARGININE N-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch.)
CC -----
CC DR EMBL; AF232716; AAF37292.1; -
CC DR EMBL; AF232717; AAF37293.1; -
CC DR MGD; MGI:107846; Hmmt112.
CC DR InterPro; IPR000051; -
CC DR Transferase; Methyltransferase; Nuclear protein; Alternative splicing.
CC FT VARSPLOC 13 30 MISSING (IN ISOFORM 2);
CC FT SEQUENCE 371 AA; 42435 MW; AEFCEF63001B1A58C CRC64;
SQ

Query Match 15.6%; Score 494.5; DB 1; Length 371;
Best Local Similarity 32.4%; Pred. No. 2,4e-29;
Matches 121; Conservative 67; Mismatches 149; Indels 37; Gaps 8;

```

```

OY 217 STMAOAEVLKSNLNRDRIIVIPGKVEEYSLP-EQVDITISEPMGYLMENERMLESTYH 275
Db 120 SSISDIAVKIVKANKLDHVVITIKGVEEVLPEKVDIITISEMGICLFESMLNTVLH 179
OY 276 AK-KYLKPSGNMPEPTIGDVLHAPFTDEQLMEQFTKANFRQPSFHGVDLSALGAANDE 334
Db 180 ARKKMLAPDGLIFPRATITVTAIED-----RQYKDKYKIHMMENYVYGDMSCKIKVAI-- 232
OY 335 YRQPVVDFTDIRILMAKSVKTYVNFLEAKEGDLHRIETI-----PKRFHMLHSGLVHG 387
Db 233 --KEPLVVDVDPKQVLV-----TNACLIKEVDIYTVKVEDLTFTSPFCLOVKKNDVYHA 283
OY 388 LAFWFDVAFISIMVWLTAPDEPLTHMYQVRCLEPSPFLFAKGDPLISGCLLIANKRQ 447
Db 284 LVAFNIEFTRCHKRTGFTSPSPSYTHMKQVTFYVMEVDYLVTKGEBIEFTIGMRPNKN 343
OY 448 SYDISIVAOVDQTG 461
Db 344 NRDLFTITDLDFKG 357

RESULT 3
HMT1_YEAST STANDARD: PRT: 348 AA.
AC P38074;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HNRNP ARGININE N-METHYLTRANSFERASE (EC 2.1.1.-) (ODP1 PROTEIN).
GN HMT1 OR ODP1 OR RMT1 OR YBR034C OR YBR0320.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=5288C;
RX MEDLINE=94378725; PubMed=8091864;
RA Smith P.H.M., de Haan M., Maat C., Grievell L.A.;
RT "The complete sequence of a 33 kb fragment on the right arm of
RT chromosome II from Saccharomyces cerevisiae reveals 16 open reading
RT frames, including ten new open reading frames, five previously
RT identified genes and a homologue of the SCO1 gene.";
RL Yeast 10:S75-S80(1994).
RN [2]
RP SEQUENCE OF 1-55 FROM N.A.
RX MEDLINE=95204349; PubMed=7896706;
RA Loubard A., Karst F., Guilloton M., Marchieau C.;
RT "sterol uptake induced by an impairment of pyridoxal phosphate
RT synthesis in Saccharomyces cerevisiae: cloning and sequencing of the
RT PDX3 gene encoding pyridoxine (pyridoxamine) phosphate oxidase.";
RL J. Bacteriol. 177:1817-1823(1995).
RN [3]
RP CHARACTERIZATION.
RA Henry M.F., Silver P.A.;
RL unpublished observations (MAR-1996).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=96218186; PubMed=8647869;
RA Gary J.D., Lin W.-J., Yang M.C., Herschmann H.R., Clarke S.;
RT "The predominant protein-arginine methyltransferase from
RT Saccharomyces cerevisiae.";
RL J. Biol. Chem. 271:12585-12594(1996).
CC -1- FUNCTION: METHYLATES ARGININES IN A VARIETY OF RNA-BINDING
CC PROTEINS. CAN CATALYZE BOTH THE MONO- AND ASYMMETRIC
CC DIMETHYLATION.
CC -1- SIMILARITY: BELONGS TO THE PROTEIN ARGININE N-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```
CC modified and this statement is not removed. Usage by end for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
  
DR EMBL; X76078; CAA53689.1;  
DR EMBL; Z35903; CAA84976.1;  
DR EMBL; X76992; CAA54296.1;  
DR PIR; S45890; S45890.  
DR SGD; S0000238; HMT1.  
KW Transferase; Methyltransferase.  
SQ SEQUENCE 348 AA; 39786 MW; 363AF61033FB4AC2 CRC64;
```

Query Match 15.3%; Score 486; DB 1; Length 348;
Best Local Similarity 36.5%; Pred. No. 9,4e-29;
Matches 120; Conservative 56; Mismatches 121; Indels 32; Gaps 9;

QY 139 SERTESSAVV-YFOFYCYLSOOQQMMMDYYRTGYORAILQHNDPDKIVLDVGCGSG 197
 :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 11 TEKRLTSSSEQHYFNSVDYHGHEMLDQTVATLTSYRANILNKKLFKDIYLVDCGCTG 70
 |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 198 ILSPFAQAQARKIAYEASTMAQHAELVKSNNTLTRIVIPGEVEVSCLP-EQVDIT 256
 ||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 71 ILSMFAAHKGAKHVIYGVMSSSIEMAKELELVNGFSDDKITTLRGKLEDVHLPEFPYDITI 130
 |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 257 SEPWCYMLENRMLESYLIAHK-KLYAPESGNMPPTIGDVHLAFTPDOLMEFTANRRY 315
 ||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 131 SEMMGYFLYESMDQTVIYARDHYLEVGGLLPPDKCSIHLAGLEP---SQYKDCKINY 185
 ||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 316 QPSFHGVDSLALRGAAVDEYEPQPVDFEDIRILMAKSVKYVNFLEAK--EGDLRIET 373
 |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 186 WQDYGFYSPE---VLPLVLPETIDYE-----RNNVNTSDKLIEFDINITYKI 232
 |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 374 P-----FKFEHMHSLGHGLAEFFDAVFIG--SIMTWISLTAPTETPTHWYOVCLEQ 424
 |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 233 SDLAFRSNEFKLKARQDMINGIWTFEDIYPAPKGRKPVEESTGPAPHYTEHWKOTFEYP 292
 |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
QY 425 SPLPAKAGDTLSGTOLLANKROSQDISI 453
 |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db 293 DDLDETGDITLEGELVSCSPNEKNRDNLNI 321
 |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:

RESULT 4
ANM4_HUMAN STANDARD; PROT; 334 AA.

AC O9NRJ2;
Dt 01-OCT-2000 (Rel. 40, Created)
Dt 01-OCT-2000 (Rel. 40, Last sequence update)
Dt 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTEIN ARGinine N-METHYLTRANSFERASE 4 (EC 2.1.1.-).
GN HRMT1L3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lorenz B.; Strom T.M.;
RT "Transcripts in human map region 12p13.3."
RL Submitted (MAR-2000) to the EMBL/Genebank/DDBJ databases.
CC -! FUNCTION: PROBABLY METHYLATES THE GUANDINO NITROGENS OF ARGINYL
CC RESIDUES IN SOME PROTEINS (BY SIMILARITY).
CC -! SIMILARITY: BELONGS TO THE PROTEIN ARGinine N-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by end for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

CC -----CC
DR EMBL, AF263539; AAP19390.1; -
DR InterPro; IPR000051; -
DR InterPro; IPR001781; -
KW Transferase; Methyltransferase.
SV
SQ SEQUENCE 334 AA; 38835 MW; C3BC4373159B224A CRC64;

Query Match 15.3%; Score 484.5; DB 1; Length 334;
Best Local Similarity 33.8%; Pred. No. 1.2e-28;
Matches 111; Conservative 65; Mismatches 127; Indels 25; Gaps 6;

OY 143 EESSAVOYFQFYGYLSQQOONMODVPRGTQORALIHQHTDEKDKIVLDVCGGSGILSPF 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9 EETSDYTFDSTAHNGHIEEMKDEKRLITLTKNSMHWKHNKVDKVDLDVGSSTGLISMF 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 203 AAAGARKITVAEASTMAOHAELVKNSNNLDRIVVIGKVEVSLP-EQVDIIISSEPMG 261
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 69 AAKAGAKKVFGECSISIDYSEKIKKANHLNDIITFPKGVEVELPEPKVDIISSEMG 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 262 YLMEERNELESTLIAK-KYLKPSGMFFPTIGDVHLAPFTDEOLYEQFTKANFRQPSFH 320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 129 YCLFESMLNVIYFARDKMKLPGGLMPFDRAALYVALED-----ROYKDEKIHMMENY 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 321 GVDLSALGAANDVEFROPVDTEPDIRILMKSVYTYNPLEAKSGDLHRIE----- 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 184 GDMTCICIDVAA---KEPLVDIVPK-----OVYTAACLKEKEDYITVTEELSPFS 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 374 PFKFMHLHSGLVHGLAFEDVAFISIMTVLSTAPTEPLTHMYVRCLOFSPLEAKAGD 433
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 233 AFCLDIQNDYVHALVTFENIEFTGCHKMGSGTAPDAFYHWKQTVFLBEDYLVVRSGE 292
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

OY 434 TLGFTCLLIANKROSYDISIVAQVDQTG 461
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 293 ELYGTISMKNPAKNVRLDFTVDLDFKG 320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
ANML_HUMAN STANDARD; PRT; 361 AA.
AC Q99873; Q99872; Q99874; Q15529; Q9NZ04; Q9NZ05; Q9NZ06;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTEIN ARGININE N-METHYLTRANSFERASE 1 (EC 2.1.1.-) (INTERFERON
DE RECEPTOR 1-BOUND PROTEIN 4).
GN HRMT1L2 OR PRMT1 OR HMT2 OR IRB4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=98207248; PubMed=9545638;
RA Scott H.S., Antonarakis S.E., Laliot M.D., Rossier C., Silver P.A.,
RA Henry M.F.;
RT "Identification and characterization of two putative human arginine
RL methyltransferases (HRMT1L1 and HRMT1L2).";
RL Genomics 48:330-340(1998).
RN [2]
RP SEQUENCE FROM N.A.
RP TISSUE=Glial cell;
RX MEDLINE=96257203; PubMed=8675017;
RA Nikawa J., Nakano H., Ohi N.;
RT "Structural and functional conservation of human and yeast HCPI genes
RT which can suppress the growth defect of the Saccharomyces cerevisiae
RL Irls mutant.";
RL Gene 171:107-111(1996).
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RP Scorillas A., Black M.H., Diamandis E.P.;
RT "Molecular characterization, mapping and tissue expression of the
RT human protein arginine N-methyltransferase 1 (PRMT1/HRMT1L2) gene.";

```

Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: METHYLATES (MONO AND ASYMMETRIC DIMETHYLATION) THE GUANIDINO NITROGENS OF ARGINYL RESIDUES PRESENT IN A GLYCINE AND ARGININE-RICH DOMAIN (CAN METHYLATE HNRNAI AND HISTONES) (BY SIMILARITY).

-1- SUBCELLULAR LOCATION: NUCLEAR.

-1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1/2 (SHOWN HERE), 2/3 AND 3/4; ARE PRODUCED BY ALTERNATIVE SPLICING.

-1- SIMILARITY: BELONGS TO THE PROTEIN ARGININE N-METHYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; Y10806; CAA71764.1; -
 DR EMBL; Y10805; CAA71763.1; -
 DR EMBL; Y10807; CAA71765.1; -
 DR EMBL; D66904; BAA11029.1; -
 DR EMBL; AF222689; AAF62893.1; -
 DR EMBL; AF222689; AAF62894.1; -
 DR EMBL; AF222689; AAF62895.1; -
 DR MIM; 602950; -
 DR transferase; Methyltransferase; Nuclear protein; Alternative splicing.
 FT VARSPLIC 1 19 MENFVATLNGMSLOPPE -> MGVSA (IN ISOFORM 2).
 FT VARSPLIC 1 19 MENFVATLNGMSLOPPE -> MGVSA (IN ISOFORM 2).
 FT VARSPLIC 1 19 MENFVATLNGMSLOPPE -> M (IN ISOFORM 3).
 FT CONFLICT 108 108 V -> E (IN REF. 3).
 FT CONFLICT 147 175 DIITSEMGTCLEYSMLNLYVARDKWL -> ASSASGWA
 FT SEQUENCE 361 AA; 41485 MW; A288969B5AF8760 CRC64;

Query Match 15.2%; Score 481.5; DB 1; Length 361;
 Best Local Similarity 31.8%; Pred. No. 2.1e-28;
 Matches 119; Conservative 68; Mismatches 150; Indels 37; Gaps 8;

QY 97 OSFIITLGNNSVLIQFATPHDFCSFYNIILKTCRGHTLERSVFSESRPESSAVOYFOFYGY 156
 DB 2 ENFVALI-ANGMSLOPPELVSCG-----QAESSEKPAEMDMTKDYFDSTYAH 49

QY 157 LEOQQNMADYVFTGYQRAILLQNHDFDKIVLDVGGSGILSFFAAQAGRIYAVEA 216
 DB 50 FGJHEMLKDEVTLYTRNSMFHNRHLFKDKVLDVGGSGILCMFAAKAGARKYVIGIVC 109

QY 217 STMAQAEVLVKNLNLDRIVLPGRVEESLP-EQVDIIITSEPMGYMLFNERMLESYLH 275
 DB 110 SSISDVAIVKANKLDHVVTIIGKRVEEVELPEKVDIIITSEMGYCLYESMLNLYV 169

QY 276 AK-KYLKPSGNMPTIGDVLAFPTDEQLYMEQFTKANFRPSFHVDSALGGAIVDE 334
 DB 170 ARKMLAPDGLIFPDRAIYVTAIED-----ROYKDKIMWENVYGFDSCLIDVAI-- 222

QY 335 YEPQPVVDFTDIRILMAKSVKYTVNLEAKEGDLHRIE-----DFKSHMLHSGLVHG 387
 DB 223 --KEPLVDVDPKQV-----TNACLKEVDIYKAVVEDLFTSPCLQVKNNDYVHA 273

QY 388 LAWFEDVAPLGSIMYVLSAPTEPLETHYQVRCLEFSPFLAKAGDLSGLCLLIANKRQ 447
 DB 274 LVAVFNIIEFRCRKRGTSPSPSPYHMKQTVFYMEDYLVTKGIEIFGTIGMRPAKN 333

QY 448 SYDISIAOVDTG 461
 DB 334 NRDLPTDIDDFG 347

RESULT 6
 ANM2_MOUSE STANDARD; PRT; 448 AA.
 ID ANM2_MOUSE

AC 09R144;
 DT 01-OCT-2000 (Rel. 40; Created)
 DT 01-OCT-2000 (Rel. 40; Last sequence update)
 DT 01-OCT-2000 (Rel. 40; Last annotation update)
 DE PROTEIN ARGININE N-METHYLTRANSFERASE 2 (EC 2.1.1.-).
 GN HMTLLI OR PMT2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tavba L.O., Kvasna S.M., Skripkina I.Y., Anoprienko O.V., Slavov D., Tassone F., Rynditch A.V., Gardiner K.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROBABLY METHYLATES THE GUANIDINO NITROGENS OF ARGINYL RESIDUES IN SOME PROTEINS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PROTEIN ARGININE N-METHYLTRANSFERASE FAMILY.

CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF169620; AAD4847.1; -
 DR MGD; MGI:131652; Hmtll1.
 DR HSDP; P06241; ISHF.
 DR InterPro; IPR000051; -
 DR InterPro; IPR001452; -
 DR PRINTS; PR00452; SH3DOMAIN.
 DR PROSITE; PS50002; SH3; 1.
 DR transferase; Methyltransferase; SH3 domain.
 FT DOMAIN 42 101 SH3.
 FT SEQUENCE 448 AA; 50476 MW; 63E1F8C36C6C25E CRC64;

Query Match 15.0%; Score 475; DB 1; Length 448;
 Best Local Similarity 31.0%; Pred. No. 8.8e-28;
 Matches 138; Conservative 64; Mismatches 171; Indels 72; Gaps 14;

QY 23 GAGPCATVVFPGARLLTIGDANGELIORHAPQALREVRAGPDAGIALISHEVCYFK 82
 DB 5 GEGCSESQVVP---VLEEDVDVGCCEMQLQDQGLQLOLOPPE-----EFVAIAD 52

QY 83 CSVRETECSHVGROSFTITLGCNSVLIQFAPPHDFCSFYNIILKTCRGHTLERSVFSE 142
 DB 53 YTADEDTQLS-----FLRGEKILLRQTTADM--WMSRACCGY-IPANILGKOL 100

QY 143 EESSA-----VOYFOFYGYLSOOQNMADYVFTGYQRAILLQNHDFDKIVLDVGGSG 196
 DB 101 EYDEPEDTWOEEFDSGLIKLHLEMADQPRTKYHSVILQNNESLKDXYLIDVGGCT 160

QY 197 GILSFPAA-QAGARKIYAVEASTMAQHAELVYKSNLNLDRIVYIPGKVEVSLPEQVDII 255
 DB 161 GILSFCNHHARPAKYVAVEASDMAQHTSQLVQNGFADTIIVFOOKVEDVVLPEKVDVL 220

QY 256 ISEPMGYLLENRMLESYLAH-KYLKPSGNMPTIGDVLAFPTDEQLYMEQFTKANFR 314
 DB 221 VSENMGTCLLEFEMESLIVARDTWLKGDIITPTAALHVLPCSAEKDYHSKVLFWDNA 280

QY 315 YQSFHGVDSALGAAVDEYFROP-----VVDFTDIRILMAKSVKYTV 358
 DB 281 YE-----FNLSTAKSLAIKEFFSRKSNHLKPEDCLSEPCIIILDHRTVQ-----V 328

QY 359 NFLAKESGDLHRIEIPKFFHMLHSGLVGLAFWEDVAF-----IGSIMYVLSAPTEPLE 414
 DB 329 PDLEMRGEL-----RFDIQKAGTLHGFTAMFESVYFQSLSEGOPOOV-VSTGLHPPT 380

OY 415 HWYVRCLEFSPLEFAKAGDTLSTGC 439
 DB 381 HMKOTLEFMDPPVPHVGTGVHGF 405

RESULT 7
 ANM3_HUMAN STANDARD; PRT; 512 AA.

AC 060678;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROTEIN ARGININE N-METHYLTRANSFERASE 3 (EC 2.1.1.-) (FRAGMENT).
 GN PRMT3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98307932; PubMed=9642256;
 RA Tang J., Gary J.D., Clarke S., Herschman H.R.;
 RT PRMT 3, a type I protein arginine N-methyltransferase that differs
 RT from PRMT1 in its oligomerization, subcellular localization,
 RT substrate specificity and regulation.*
 RL J. Biol. Chem. 273:16935-16945(1998).
 CC -1- FUNCTION: METHYLATES (MONO AND ASYMMETRIC DIMETHYLATION) THE
 CC GUANIDINO NITROGENS OF ARGINYL RESIDUES IN SOME PROTEINS.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE PROTEIN ARGININE N-METHYLTRANSFERASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: AF059531; AAC39837.1;
 CC MIM: 603190;
 DR InterPro: IPR000051;
 DR InterPro: IPR000822;
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; UNKNOWN_1.
 KM Transferrase: Methyltransferase; zinc-finger.
 FT NON_TER 1 1
 FT ZN_FING 29 52 C2H2-TYPE.
 SQ SEQUENCE 512 AA; 58098 MW; 260DC4EB25162A18 CRC64;

Query Match 15.0%; Score 474.5; DB 1; Length 512;
 Best Local Similarity 33.4%; Pred. No. 1.2e-27;
 Matches 121; Conservative 67; Mismatches 137; Indels 37; Gaps 12;

OY 111 OFATPHDFCSFYILKTCRGHTLERSYFSESTESSAVOYFOFYGLYSOOONMADYVRT 170
 DB 169 QFA--QDFV-MHTDVRCSSST---SVIADLODEDEGV-FSSYGHGIGHHEMLKDKIRT 221
 OY 171 GTYQRAIILNHNDFKDKIVLDVCGSGIISFFAAQAGARKIYAEASTMAQHAETVLYKSN 230
 DB 222 ESYRDELYQNHIFKDKVVDVCGSGTILSMFAKAKAKVIGVDSIILYQANDIIRLN 281
 OY 231 NLUDRIYVPGKYEVSLEP-EQVDIIISEPMGYLMEHMEIESTLAK-KYLKSGNMFP 288
 DB 282 KLEDTITLLKGLKEEYVHLPEKVDVILISEWNGFLLFESMDSVLYAKNKLARGGSVYP 341
 OY 289 TIGDVHIAFTPEQLYMEQFTKANFRYQSPFHGVDSLALRGAAVDEYFROPVVDPIRI 348
 DB 342 DICTISLVANSVDNKNHADRIA-----FWDDVYGGKMSCKMKAVIPE-----AYVVDLPKT 392
 OY 349 LMAK--SVKTYVNLKAEKEDLARIET-----PFRFHMILHGLVHGLAFWDPVAETGS 399

DB 393 LISEPGGKIH-----DCHTTSISDLERSSDTLKTRTSMCTALAGYEDIVEKN 443
 OY 400 IMT-VMLSTAPTEPLTHWYOVRLCFOSPLFAKAGDTLSTGCLLIANKRPOSVDISIYAOVD 458
 DB 444 CHNRVFEFTGPSTKTHMKQIVLEKPEPSVAGALKGVYHKKKPRSLTVILTLN 503

OY 459 QT 460
 DB 504 NS 505

RESULT 8
 ANM1_SCHPO STANDARD; PRT; 339 AA.

AC 090RX7;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROBABLE PROTEIN ARGININE N-METHYLTRANSFERASE (EC 2.1.1.-).
 GN SPAC890.07C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_Taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Saunders D., Harris D., McDougall R.C., Rajadream M.A., Barrell B.G.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROBABLY METHYLATES THE GUANIDINO NITROGENS OF ARGINYL
 CC RESIDUES IN SOME PROTEINS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PROTEIN ARGININE N-METHYLTRANSFERASE
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: AL133498; CAB63498.1;
 DR InterPro: IPR000051;
 DR InterPro: IPR000051;
 KM Hypothetical protein; Transferrase; Methyltransferase.
 SQ SEQUENCE 339 AA; 38915 MW; 12B4988ED5E17614 CRC64;

Query Match 14.9%; Score 472.5; DB 1; Length 339;
 Best Local Similarity 36.9%; Pred. No. 9.1e-28;
 Matches 114; Conservative 49; Mismatches 131; Indels 15; Gaps 7;

OY 150 YFOFYGLYSOOONMADYVRTGTQRAIILNHNDFKDKIVLDVCGSGIISFFAAQAGAR 209
 DB 19 YEDSYSHGIGHHEMLKDVDRITSLDALNQNPHLRDKIVLDVCGSGTILSMFCANAGAK 78
 OY 210 KIYAVEASTMAQHAETVLYKSNLIDRIYVPGKYEVSLEP-EQVDIIISEPMGYLMEHMEI 268
 DB 79 HVGDMSEIITHKAQVIEVNKLSRITLIGKMEIQLPEKVDIYSEWNGFLLYES 138
 OY 269 MLESTLHAK-KYLKSGNMFPITGVNHLAPFDEQLYMEQFTKANFRYQSPFHGVLSAL 327
 DB 139 MLDVYLVARDKRIAPDGLFPDRAOQLAIEDADYKSEK-----IGFDDVYGFDPSP 193
 OY 328 RGAADVDEYFROPVVDTP-IRILAKSVKTYVNLKAEKEDLARIETPKFHMILHGLVH 386
 DB 194 K---KDVAKREPLVDVIRIANTSCVILDDLTVKKED-ARSSPEIRIATRNDPVA 248
 OY 387 GLAFWEDVAFIGSINTVLSAPTEPLTHWYOVRLCFOSPLFAKAGDTLSG---TCLLIAN 444
 DB 249 AFLAFWEDIEFACHKPIKSTGPFSTRYTHMKQIVLEKPEPSVAGALKGVYHKKKPRSLTV 308

OY 445 KROSYDISI 453
DB 309 NHRELDIDI 317

RESULT 9

ANM3_RAT STANDARD; PRT; 528 AA.
AC 070467;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTEIN ARGININE N-METHYLTRANSFERASE 3 (EC 2.1.1.-).
GN PRMT3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98307932; PubMed=9642256;
RA Tang J., Gary J.D., Clarke S., Herschman H.R.;
RT "PRMT 3, a type I protein arginine N-methyltransferase that differs
RT from PRMT1 in its oligomerization, subcellular localization,
RT substrate specificity, and regulation."
RL J. Biol. Chem. 273:16935-16945(1998).
CC -1- FUNCTION: METHYLATES (MONO AND ASYMMETRIC DIMETHYLATION) THE
CC GUANIDINO NITROGENS OF ARGINYL RESIDUES IN SOME PROTEINS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE PROTEIN ARGININE N-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF059530; AAC40158.1; -.
CC InterPro: IPR000051; -.
CC InterPro: IPR000822; -.
CC PROSITE: PS00028; ZINC_FINGER_C2H2; UNKNOWN_1.
CC TRANSFERASE; Methyltransferase; Zinc-finger.
CC ZN_FING 46 69 C2H2-TYPE.
CC FT
CC ZN_SEQUENCE 528 AA; 59419 MW; B25D627902594B39 CRC64;
CC SQ

Query Match 14.9%; Score 472; DB 1; Length 528;
Best Local Similarity 34.2%; Pred. No. 1.9e-27;
Matches 123; Conservative 63; Mismatches 140; Indels 34; Gaps 11;

OY 111 OPATPHDFSEFVILKTCRGHTLERSVSESTRESSAVOFEFGYLSOOONMMQDVVRI 170
DB 186 QRA--QDFWNVVY-RTCSST---TIALQDEDEGV-YFSSYGHGHEEMAKDKVRI 237
OY 171 GYVRAAILNHTDFKDKIVLDVCGSGSILSFPAAGARRIYVEASTMAQHAELVKN 230
DB 238 ESRDITVYGNPHFKDKVLDVCGSGSILSMFAKAKKIVINVDSEILLYQANDIIRLN 297
OY 231 NLTDRIVIPGKVEVSLP-EQVDIIISSEPMGYLLENRMLESYLAK-KYLKPSGMFP 288
DB 298 KLEDTIVLJKKIEEVSLEPEKVDVSISEMGYFLFESMLDVLVYAKSKYLAAGSGSV 357
OY 289 TIGDVALAFTDQLYMEQCTKANFRQPSFHGVDLSALGAANDVEFRQPVDTPEIRI 348
DB 358 DICTISLVAVSDSKHADRIA-----FWDVYGFNMSCMKKAVIPE---AVEEVVDHKT 408
OY 349 LMAKSVKYVNFLEAKEGDLRIE-----PFKFMHLSGVLHGLAFWDFVAFISIM 401
DB 409 LISDP-----CDIKHIDCHTISISDLERSSDFTLTRTTAMCTAVAGFDIIFEKNC 461

OY 402 T-VMLSNAPTEPLTHWQVRCFLPSPLFAKAGDITSGTCLLIANKROSYDISIAVOVD 460
DB 462 NRYVSTGPOSTKTHMKQITFLLEKPEPVKAGELKKKITYVHKKKDPSRLIVLTLSNS 521

RESULT 10

ANM2_HUMAN STANDARD; PRT; 433 AA.
AC P53345; P78350; Q9JWC2;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTEIN ARGININE N-METHYLTRANSFERASE 2 (EC 2.1.1.-).
GN HRMTL1 OR PRMT2 OR HMT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98207248; PubMed=9545638;
RA Scott H.S., Antonarakis S.E., Lalioti M.D., Rossier C., Silver P.A.,
RA Henry M.F.;
RT "Identification and characterization of two putative human arginine
RT methyltransferases (HRMTL1 and HRMTL2)."
RL Genomics 48:330-340(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97341120; PubMed=9196002;
RA Katsanis N., Vaspo M.-L., Fisher E.M.C.;
RT "Identification and mapping of a novel human gene, HRMTL1,
RT homologous to the rat protein arginine N-methyltransferase 1 (PRMT1)
RT gene."
RL Humm. Genome 8:526-529(1997).
RN [3]
RP SEQUENCE OF 147-277 FROM N.A.
RA Auffray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,
RA Lebrach H., Poustka A., Lundeberg J.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLY METHYLATES THE GUANIDINO NITROGENS OF ARGINYL
CC RESIDUES IN SOME PROTEINS.
CC -1- TISSUE SPECIFICITY: UBIOUITOUS.
CC -1- SIMILARITY: BELONGS TO THE PROTEIN ARGININE N-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X99209; CA67599.1; -.
CC EMBL: U080213; AAB48437.1; -.
CC EMBL: AL109794; CAB52454.1; -.
CC DR HSSP: P06241; 1A2G.
CC DR MIM: 601961; -.
CC DR InterPro: IPR001452; -.
CC DR Pfam: PF00018; SH3_1.
CC DR PRINTS: PR00452; SH3DOMAIN.
CC DR PROSITE: PS50002; SH3_1.
CC FT TRANSFERASE; Methyltransferase; SH3 domain.
CC FT DOMAIN 30 89 SH3.
CC FT CONFICT 256 260 KDYSR->RIYA (IN REF. 2).
CC SQ SEQUENCE 433 AA; 49041 MW; 6DEB0350C15CDAF CRC64;
CC SQ

Query Match 14.6%; Score 461; DB 1; Length 433;
Best Local Similarity 34.7%; Pred. No. 9.1e-27;
Matches 119; Conservative 53; Mismatches 123; Indels 48; Gaps 10;

FT DOMAIN 238 329 COIL 1B.
 FT DOMAIN 330 353 LINKER 12.
 FT DOMAIN 354 492 COIL 2.
 FT MOD_RES 57 57 PHOSPHORYLATION (BY SIMILARITY).
 SO SEQUENCE 638 AA; 65871 MW; 98743AB872076AF CRC64;

Query Match 3.9%; Score 122.5; DB 1; Length 638;
 Best Local Similarity 21.3%; Pred. No. 0.18; Indels 175; Gaps 31;
 Matches 132; Conservative 74; Mismatches 240;

10 GPG--AGSAGVAGGAGGATVSEPCARLLTGDANGELIORHAEQCAL--RLVPRAGP 65
 130 GPGVFGGPGSGGPGGPGP-----GGFPGG-----IQEVVNSLQPLVNEIDP 174
 66 DAAGIALYSHEDVGVFCVSVRETECSRVGROSFITTLGNSVLIQFATPHDFCSFYNL 125
 175 QIGGVKAQERQ-----IKTL--NNKFASFIDKVFLEQOKV 210
 126 KTCGHFLERVSERTESSAVQYFQY-----GYLSQQQNMADYVKT 170
 211 LETWELLQOO--TTGSGPSSLEPCFESYISFLCKQLDLSLGERNLTGELKSMODLVE- 267
 171 GTYGRATLQNTDKDKTVLDVGGSGSLSPFAQAQARKIYAVASMAQAEVLVYSN 230
 268 -----DKKKIEDINKRTAAENEF--VGLKK--DVDAAFM--NKVELQAKVD 309
 231 NLADRIVIVPEKVEEVSLEQVDIIISEPMGYMLFNEHMLSEYLHAKK--YLKPSGNMF 287
 310 SLIDVEVSFL-----RTLYEMELSQMSHASTPSVLSMDNRC 347
 288 PTIDVHLAPTD--EQLYMEQFTANFRYQPSF-----HGVDSALRGAAVDERFR 337
 348 LDLSIIAEVTOYEIEIQORSKEAEALYQTKLQTLTAGHGDLENTK--SEIMEENR 406
 338 QPVVDFTDIRLAKSVKYTYNPLEAKE--GDLHRIEIPFKFMHLSGLVHGAEFVDAF 396
 407 MIQRLAREIENKQANMLQTAIAEORGEMALKDANAKLODLOTAQOKAD--DLAR 463
 397 I---GSIMTWLSTAPTEPLTHWYQVRCLPFSPLEAKAGTLSTGCLLIANKROSVDIS 452
 464 LLROYELMANKLA--LDEIATY---RKLIEGECRMSGEGSAYC-----IS 507
 453 IVAQVDOTGSKSNLLDKNPFERYTGTPPPPGSHYTSBSENMNTGYTNLSSGVAV 512
 508 VVSNVTSTSGSSGSSRGV-----FGVSGSGSGG-----YKGGSSSSSSSGYGV 551
 513 A-GMPTAV--DLSSVIAG-----SSVGHNNLIPLANTGIVNHTSRMGSIMSTGIY 561
 552 SGGSGSGTGVSGSTGGRSGSGSYOSSSSGR---LGGAGSISVSHSGMS--SSGSI 605
 562 QGSSGA---OGGSGSSSAHYA 579
 606 QTSGSGYKSGGSGSTIRFS 626

RESULT 13
 YBKL YEAST STANDARD; PRT; 1077 AA.
 AC P38167;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE HYPOTHEICAL 119.3 KDA PROTEIN IN SPT1 INTERGENIC REGION.
 GN YB101C OR YB10814.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RA [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8288C;
 RX MEDLINE=96076635; PubMed=7502586;

RA Obermayer B., Gassenhuber J., Piravandi E., Dondoy H.;
 RT "Sequence analysis of a 78.6 kb segment of the left end of
 RT Saccharomyces cerevisiae chromosome II.";
 RL Yeast 11:1103-1112(1995).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL, X79489; CAAS5995.1;
 DR EMBL, Z35862; CAB39760.1;
 DR PIR, S45395; S45395.
 DR SGD, S0000197; ECM21.
 KM Hypothetical protein.
 SO SEQUENCE 1077 AA; 119344 MW; 7F4608682B8081D3 CRC64;

Query Match 3.8%; Score 120.5; DB 1; Length 1077;
 Best Local Similarity 23.0%; Pred. No. 0.53;
 Matches 73; Conservative 47; Mismatches 119; Indels 79; Gaps 16;

316 QPSFHGVDLSALRGAAVDEFFROPVYDTDIRLAKSVKYTYNPLEAKEGDLHRIEIPF 375
 733 RPSVIGF--LSGHRKSGSHENKRPVYDP-----KFHOTIRKSNHG-----LPV 774
 376 KPHMLHSGVHL--AFEDVAFTGIMTWLSTAPTEP-----LTHWYQVRCFOSPL 427
 775 KHTRLNTPKRLYLDLSLFSNVYCRHKLEIMRISKPECESKLRH--YEV--LDTPI 831
 428 FAKAGDTLSTGCLLIANKROSVDISIVAQVDOTGSKSNLLDKNPFERYTGTPSP-- 485
 832 FLVSQCNQNGNML-----PTYDMATM-----EGKKNQVPLSNQSPF--GNCPPTPTF 879
 486 -----PGSHYTSBSENMNTGSTY-----NLSSGVAVAGPTAVDLSSVIAGGS 529
 880 EBAISVPASPIVSPMGS--NIMASYDPDLISIOQLMISRTTYSVSGHPDTSDDAGV----- 933
 530 SVGHNNLIPLANTGIVNHTSRMGSIMSTGIQSGSAGGSGSSSAHYAVNQFTMGCP 589
 934 -----PNVNRNSTSNANANGSISNSAFVSGNSG--GVARARATSVNDRSRFNNLDK 984
 590 AISMASPM-----SIPTN 602
 985 LLSTPSPVNRSHNSPTN 1002

RESULT 14
 YB3 YEAST STANDARD; PRT; 827 AA.
 AC P38374;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHEICAL 95.2 KDA PROTEIN IN AGP2-CKS1 INTERGENIC REGION.
 GN YBR133C OR YBR1008.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RA [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8288C;
 RX MEDLINE=94378717; PubMed=8091856;
 RA Becam A.-M., Cullin C., Grzybowski E., Lacroite F., Nasr F.,
 RA Ozler-Kalogeropoulos O., Paluch A., Slonimski P.P., Zagulski M.,
 RA Herbert C.J.;
 RT "The sequence of 29.7 kb from the right arm of chromosome II reveals
 RT 13 complete open reading frames, of which ten correspond to new
 RT genes."

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen, Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 09:57:41 ; Search time 37.19 Seconds

Title: US-09-464-377-2

Sequence: 1 MAAATAATGCGAGSAGVAG.....PAISMASPMSIPTNTMHYGS 608

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database :

- ```

1: sp. archaea: *
2: sp. bacteria: *
3: sp. fungi: *
4: sp. human: *
5: sp. invertebrate: *
6: sp. mammal: *
7: sp. mhc: *
8: sp. organelle: *
9: sp. phage: *
10: sp. plant: *
11: sp. rodent: *
12: sp. unclassified: *
13: sp. vertebrate: *
14: sp. virus: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID      | Description         |
|------------|-------|-------------|--------|----|---------|---------------------|
| 1          | 3168  | 100.0       | 608    | 11 | Q9WVG6  | Q9WVG6 mus musculus |
| 2          | 1409  | 44.5        | 530    | 5  | Q9VH48  | Q9VH48 drosophila   |
| 3          | 895.5 | 28.3        | 388    | 10 | Q9M906  | Q9M906 arabidopsis  |
| 4          | 864.5 | 27.3        | 577    | 10 | Q9FT88  | Q9FT88 arabidopsis  |
| 5          | 498.5 | 15.7        | 390    | 10 | Q9S094  | Q9S094 arabidopsis  |
| 6          | 496.5 | 15.7        | 376    | 10 | Q81813  | Q81813 arabidopsis  |
| 7          | 478   | 15.1        | 376    | 5  | Q9VGW7  | Q9VGW7 drosophila   |
| 8          | 477.5 | 15.1        | 383    | 10 | Q9MA15  | Q9MA15 arabidopsis  |
| 9          | 470   | 14.8        | 345    | 5  | Q9N8C5  | Q9N8C5 trypanosoma  |
| 10         | 467.5 | 14.8        | 366    | 10 | Q882210 | Q882210 arabidopsis |
| 11         | 453.5 | 14.3        | 355    | 5  | Q9VQX9  | Q9VQX9 drosophila   |
| 12         | 452.5 | 14.3        | 316    | 4  | Q9NVR8  | Q9NVR8 homo sapien  |
| 13         | 452.5 | 14.3        | 380    | 10 | Q9SNQ2  | Q9SNQ2 oryza sativ  |
| 14         | 449.5 | 14.2        | 348    | 5  | Q9U2X0  | Q9U2X0 caenorhabdi  |
| 15         | 443   | 14.0        | 399    | 10 | Q9LJZ9  | Q9LJZ9 arabidopsis  |
| 16         | 426   | 13.4        | 341    | 5  | Q9VFP8  | Q9VFP8 drosophila   |
| 17         | 394.5 | 12.5        | 602    | 10 | Q9LHM9  | Q9LHM9 arabidopsis  |
| 18         | 390   | 12.3        | 324    | 5  | Q9VKL4  | Q9VKL4 drosophila   |
| 19         | 384.5 | 12.1        | 516    | 5  | Q9VFB3  | Q9VFB3 drosophila   |

|    |       |      |      |    |        |                     |
|----|-------|------|------|----|--------|---------------------|
| 20 | 374   | 11.8 | 543  | 3  | 013648 | schizosacch         |
| 21 | 349   | 11.0 | 521  | 4  | Q9P6B1 | Q9P6B1 neurospora   |
| 22 | 271.5 | 8.6  | 284  | 4  | Q9J781 | Q9J781 homo sapien  |
| 23 | 231   | 7.3  | 5060 | 2  | Q9K5M1 | Q9K5M1 anabaena sp  |
| 24 | 223   | 7.0  | 405  | 5  | Q9NK02 | Q9NK02 leishmania   |
| 25 | 221   | 7.0  | 313  | 5  | Q9VFP9 | Q9VFP9 dtrosophila  |
| 26 | 211   | 6.7  | 161  | 4  | Q9H3E6 | Q9H3E6 dtrosophila  |
| 27 | 185.5 | 5.9  | 692  | 4  | Q9NVM4 | Q9NVM4 homo sapien  |
| 28 | 178.5 | 5.6  | 167  | 4  | Q9H5R7 | Q9H5R7 homo sapien  |
| 29 | 172   | 5.4  | 630  | 5  | Q9XMA2 | Q9XMA2 caenorhabdi  |
| 30 | 170.5 | 5.4  | 406  | 5  | Q9W1V1 | Q9W1V1 dtrosophila  |
| 31 | 170.5 | 5.4  | 637  | 4  | Q14744 | Q14744 homo sapien  |
| 32 | 168.5 | 5.3  | 633  | 4  | Q9UKH1 | Q9UKH1 homo sapien  |
| 33 | 163   | 5.1  | 670  | 10 | Q9M0J0 | Q9M0J0 arabidopsi   |
| 34 | 162.5 | 5.1  | 631  | 11 | Q9OZS9 | Q9OZS9 mus muscoli  |
| 35 | 146   | 4.6  | 259  | 1  | Q58847 | Q58847 methanococc  |
| 36 | 138   | 4.4  | 680  | 5  | Q02325 | Q02325 caenorhabdi  |
| 37 | 134   | 4.2  | 613  | 4  | Q9H9J0 | Q9H9J0 homo sapien  |
| 38 | 125.5 | 4.0  | 244  | 4  | Q9EZC1 | Q9EZC1 staphylococ  |
| 39 | 125.5 | 4.0  | 589  | 10 | Q9LW00 | Q9LW00 cryza salici |
| 40 | 122   | 3.9  | 928  | 11 | Q08963 | Q08963 rattus norv  |
| 41 | 120.5 | 3.8  | 208  | 1  | Q9UXW9 | Q9UXW9 pyrococcus   |
| 42 | 118   | 3.7  | 245  | 2  | Q9KSJ9 | Q9KSJ9 vibrio chol  |
| 43 | 117.5 | 3.7  | 1320 | 5  | Q9W001 | Q9W001 dtrosophila  |
| 44 | 116.5 | 3.7  | 610  | 5  | Q9U6Y9 | Q9U6Y9 dtrosophila  |
| 45 | 116   | 3.7  | 1844 | 5  | Q9V957 | Q9V957 dtrosophila  |

## ALIGNMENTS

|        |                                                                |
|--------|----------------------------------------------------------------|
| RESULT | 1                                                              |
| Q9WVG6 | PRELIMINARY; PRT; 608 AA.                                      |
| ID     | Q9WVG6                                                         |
| AC     | Q9WVG6;                                                        |
| DT     | 01-NOV-1999 (TREMBLrel. 12, Created)                           |
| DI     | 01-NOV-1999 (TREMBLrel. 12, last sequence update)              |
| DR     | 01-JUN-2000 (TREMBLrel. 14, last annotation update)            |
| DE     | PROTEIN ARGININE METHYLTRANSFERASE.                            |
| GN     | CARM1.                                                         |
| OS     | Mus musculus (mouse).                                          |
| OC     | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;        |
| OX     | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.      |
| RN     | NCBI_TaxID=10090;                                              |
| RP     | [1]                                                            |
| RX     | SEQUENCE FROM N.A.                                             |
| RA     | MEDLINE=99316081; PubMed=10381882;                             |
| RA     | Chen D., Ma H., Hong H., Koh S.S., Huang S., Schurter B.T.,    |
| RA     | Aswad D.W., Stallcup M.R.;                                     |
| RT     | "Regulation of transcription by a protein methyltransferase."; |
| RL     | Science 284:2174-2177(1999).                                   |
| DR     | EMBL; AF117887; AAD41265.1; -                                  |
| DR     | InterPro; IPR000051; -                                         |
| KW     | Transferase; Methyltransferase.                                |
| SQ     | SEQUENCE 608 AA; 65823 MW; 7765F56378269C79 CRC64;             |

  

|                           |                                        |
|---------------------------|----------------------------------------|
| Query Match               | 100.0%; Score 3168; DB 11; Length 608; |
| Best Local Similarity     | 100.0%; Pred. No. 9,4e-243;            |
| Matches 608; Conservative | 0; Mismatches 0; Indels 0; Gaps 0      |

  

|    |     |                                                              |     |
|----|-----|--------------------------------------------------------------|-----|
| Dy | 1   | MAAAATAATGVRAGSAGAVAGPGAGAPCATVSPFGARLLITGDANGELIORHAEQALRIE | 60  |
| Dy | 1   | MAAAATAATGVRAGSAGAVAGPGAGAPCAIVSPFGARLLITGDANGELIORHAEQALRIE | 60  |
| Dy | 61  | VKAGDAGIALYSHEDVCVFCKSVSRRECSRVGRNSEFTILGCNSVLIOFATPHDFCS    | 120 |
| Dy | 61  | VKAGDAGIALYSHEDVCVFCKSVSRRECSRVGRNSEFTILGCNSVLIOFATPHDFCS    | 120 |
| Dy | 121 | FYNLIKTRGHLEHSVSRSERIESSAVOYFOFYGLISQQNMMDYVNTGYORAILIION    | 180 |
| Dy | 121 | FYNLIKTRGHLEHSVSRSERIESSAVOYFOFYGLISQQNMMDYVNTGYORAILIION    | 180 |

```

QY 181 HTDEKDIYLVGGSGITLSPFAAQAQARKIYAVEASTMAOAEVLVKSNNLTDRIVIP 240
DB 181 HTDEKDIYLVGGSGITLSPFAAQAQARKIYAVEASTMAOAEVLVKSNNLTDRIVIP 240
QY 241 GKVEEVSILPEQVDIILISEPMGYMLFENEMESYLHAKKYLKPCGNMPTIGDVHLAPFD 300
DB 241 GKVEEVSILPEQVDIILISEPMGYMLFENEMESYLHAKKYLKPCGNMPTIGDVHLAPFD 300
QY 301 EQLYMEQFTKANFRYQSPFHGVDSALFGAAVDEYFROPVDTFDIRILMAKSVKYTVNF 360
DB 301 EQLYMEQFTKANFRYQSPFHGVDSALFGAAVDEYFROPVDTFDIRILMAKSVKYTVNF 360
QY 361 LEAKEGDHRHREIFPKFMLHSGLVHGLAFMDVAFISIMTWMLSTAPTEPLTHWYVR 420
DB 361 LEAKEGDHRHREIFPKFMLHSGLVHGLAFMDVAFISIMTWMLSTAPTEPLTHWYVR 420
QY 421 CLFQSPLEFAKAGDTLSTGCLLIANKRQSYDYSIAVQVDTGSKSNLLDLKNPFRTYGT 480
DB 421 CLFQSPLEFAKAGDTLSTGCLLIANKRQSYDYSIAVQVDTGSKSNLLDLKNPFRTYGT 480
QY 481 TPSPPPSGHITSPEENMANTSTYLSGVAVAGMPTAYDLSVYAGSSVGHNNLPLA 540
DB 481 TPSPPPSGHITSPEENMANTSTYLSGVAVAGMPTAYDLSVYAGSSVGHNNLPLA 540
QY 541 NTGIVNHTSHRMSGIMSTGYVQSSGAGGSSSAHYAVNNQFTMGPAISMASPMSTP 600
DB 541 NTGIVNHTSHRMSGIMSTGYVQSSGAGGSSSAHYAVNNQFTMGPAISMASPMSTP 600
QY 601 TMTMHYGS 608
DB 601 TMTMHYGS 608

RESULT 2
QYVH48
AC QYVH48 PRELIMINARY; PRT; 530 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CG5358 PROTEIN.
GN CG5358.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN=BERKELEY.
RA MEDLINE-20196006; PubMed-10731132.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
RA Abilaj J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Houson K.A., Howland T.J., Wei M.-H., Ibeagwu C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Klumel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```

```

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Metkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Ventler E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Welnslock G.M., Welnsenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhan M., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003685; AAF54471.1;
DR Flybase; FBgn0037770; CG5358.
DR InterPro; IPR000051;
SQ SEQUENCE 530 AA; 59727 MW; 61590959D6B91EEA CRC64;

Query Match 44.5%; Score 1409; DB 5; Length 530;
Best Local Similarity 59.3%; Pred. No. 2,4e-103;
Matches 264; Conservative 74; Mismatches 99; Indels 8; Gaps 2;

QY 85 VSRETECSVVGQSFITITGCSNVLIOFATPHDFGCFYILKTRCHTLEKRSYFSRTTE 144
DB 80 IADPDAQOMGRSVAVSLDADNLVRFASDQDLFRFVENVK-HLRPKSVFSQRTTE 138
QY 145 SSAVGYFQFYGLTSGQNNMODYVRTGYQRAILQNHTEFKRIYDVAGSGSILFFFA 204
DB 139 SSASQFQFYGLTSGQNNMODYVRTGYQRAILQNHTEFKRIYDVAGSGSILFFFA 198
QY 205 QAGARKIYAVEASTMAOAEVLVKSNNLTDRIVIPGKVEEVSILPEQVDIILISEPMGYML 264
DB 199 QAGARKIYAVEASTMAOAEVLVKSNNLTDRIVIPGKVEEVSILPEQVDIILISEPMGYML 258
QY 265 FNERMESTLHAKKYLKPCGNMPTIGDVHLAPFDQOLYMEQFTKANFRYQSPFHGVDS 324
DB 259 YNERMESTLHAKKYLKPCGNMPTIGDVHLAPFDQOLYMEQFTKANFRYQSPFHGVDS 318
QY 325 SALFGAAVDEYFROPVDTFDIRILMAKSVKYTVNFLEAKEGDHRHREIFPKFMLHSG 384
DB 319 TILHKEGMEKTEFROPVDTFDIRILMAKSVKYTVNFLEAKEGDHRHREIFPKFMLHSG 378
QY 385 VGLAFMDVAFISIMTWMLSTAPTEPLTHWYVRCLFQSPLEFAKAGDTLSTGCLLIAN 444
DB 379 VGLAFMDVAFISIMTWMLSTAPTEPLTHWYVRCLFQSPLEFAKAGDTLSTGCLLIAN 438
QY 445 KROSVDISIAVQVDTGSKSNLLDLKNPFRTYGTTPSPPPSGHITSPEENMANTSTYLS 498
DB 439 KROSVDISIAVQVDTGSKSNLLDLKNPFRTYGTTPSPPPSGHITSPEENMANTSTYLS 498
QY 499 -NTGIVNHTSHRMSGIMSTGYVQSSGAGGSSSAHYAVNNQFTMGPAISMASPMSTP 522
DB 499 GSRNSSMLNGISVNGIGEGMDIT 523

RESULT 3
QYVH06
AC QYVH06 PRELIMINARY; PRT; 388 AA.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PUTATIVE ARGININE METHYLTRANSFERASE.
GN F17A9.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicot; Rosidae; eucotids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

```



[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-CV. COLUMBIA;  
 RA Lin X., Kaul S., Town C.D., Beato M., Creasy T.H., Haas B., Wu D.,  
 RA Romling C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,  
 RA Bowman C.L., White O., Niernan W.C., Fraser C.M.;  
 RT "Arabidopsis thaliana chromosome III BAC F179 genomic sequence."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC016827; AF26997.1; -  
 DR Interpro: IPR000051; -  
 KW Transferase; Methyltransferase.  
 SQ SEQUENCE 388 AA; 43851 MW; 031F8360B484F656 CRC64;

Query Match 28.3%; Score 895.5; DB 10; Length 388;  
 Best Local Similarity 49.7%; Pred. No. 8e-63;  
 Matches 175; Conservative 64; Mismatches 98; Indels 15; Gaps 5;

QY 150 YFQYGYLSQOQNMMDYVGTGYQRAILLQNHDFKQKIVLDVCCSGSIIISFAAQAQAR 209  
 DB 2 YFHYGQLHQQNMLODYVGTGYAAVAMENHSDAGRVAVDVGAGSIIISFAAQAQAR 61  
 QY 210 KIYAVEASTMAQHAELVYKSNL-TDRIYVPGKVEVSLPEQVDITISEPMGYLFPNER 268  
 DB 62 HYAVASEMAEYAKRLIAGNPLFADRIITYKGVEDIELEPKADIIISEPMGYLFPNER 121  
 QY 269 MLESYLHAK-KYLKPSGNMPTIGDVLAFPTDEQLYMEQFTKANFRYOPSFGVDLSAL 327  
 DB 122 MLESYIARDFRFPKGMPTVGRIMAFSPDEFLIEAMANKAMFQOQNYGVDLPL 181  
 QY 328 RGAANDEROPVYDFDIRILMAKSVKTYVNLKAEGLHRLIEIPKTHMLSGLVHG 387  
 DB 182 YGSAHGYSQPVDAFDPRLVASFHMDIDTQMKEDFEYEDIDPKFTASACTMHG 241  
 QY 388 LAFMFPAVFGSINTVWLSTAPTEPLTHWYOVCLFQSPLEFAKAGDPLSGCIIILANKRQ 447  
 DB 242 LACMFVPLDGSIVQKRLTTPAGAPTHMVIQKCVLSQPIYVNAQGITRLHILIASAQ 301  
 QY 448 SY--DISIAQVVDOTGSKSSNLT-----DLKNPFRYGTTPSP-----PP 486  
 DB 302 SYTIDTLTKMKGPGASGGILOSSTCKFDLKEPYRMSQPAVPAQEP 353

RESULT 4  
 Q9F168 PRELIMINARY; PRT; 577 AA.  
 AC Q9F168:  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE ARGININE METHYLTRANSFERASE-LIKE PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-COLUMBIA;  
 RX MEDLINE=99397451; PubMed=10470850;  
 RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,  
 RA Miyajima N., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.  
 RT Sequence features of the regions of 1,011,550 bp covered by seventeen  
 RT pl and TAC clones."  
 RL DNA Res. 6:183-195(1999).  
 DR EMBL: AB017061; BAB10326.1; -  
 KW Transferase; Methyltransferase.  
 SQ SEQUENCE 577 AA; 64650 MW; 3ED6C7DE89144CB3 CRC64;

Query Match 27.3%; Score 864.5; DB 10; Length 577;  
 Best Local Similarity 38.7%; Pred. No. 4.3e-60;

Matches 189; Conservative 72; Mismatches 139; Indels 89; Gaps 8;  
 QY 79 CVFKCVSRETEGSRVGRQSFITLGCNSVLIQFATPHDFCSFYNIILKTCR-----GH 131  
 DB 75 CVDSDGISSAKESKESFSR-----GVYIKFEDKDSKEFDCSFECKKDAVAKGS 124  
 QY 132 TL-----BRVSFSETESSAVQYQFQGYLSQOQNMMDYVGTGYQRAILLQNHDF 183  
 DB 125 ALPNGTVVANKSKFEDDIEAASAKMFHYGGQLHQQNMLODYVGTGYTSTYHAAVENNSD 184  
 QY 184 FKDRIVLDVCGSGSILSFAAQAQARKIYAVEASTMAQHAELVYKSNL-TDRIYVPGK 242  
 DB 185 FSGRVVVDVGAGSIIISFAAQAQARKIYAVEASTMAQHAELVYKSNL-TDRIYVPGK 244  
 QY 243 VEVSLEQVDITISEPMGYLFPNERMLESYLHAK-KYLKPSGNMPTIGDVLAFPTDE 301  
 DB 245 IEDIELEPKADVILISEPMGYLFPNERMLETYVIAADRLSPNGKMFVYGRIMAFPADE 304  
 QY 302 QLYMEQFTKANFRYOPSFGVDLSALGANDEROPVYDFDIRILMAKSVKTYVNL 360  
 DB 305 FLFVEMANKALFWOQOQNYGVDLPLVYSAHQGYFQPVDAFDPRLVAFSMEFVIDFT 364  
 QY 361 -----LEAEGDLHRIE 372  
 DB 365 MMYVKEPPLTDSVPRCLKIEDTLSCYFYQSSILFSEMHPNSITVYFSSQEDQFTIED 424  
 QY 373 IPEKFMHLSGLVHGLAFMPDVAEFGSINTVWLSTAPTEPLTHWYOVCLFQSPLEFAKAG 432  
 DB 425 IPLKFTASVCRIRHGLACMFVPLDGSIVQKRLTTPAGAPTHMVIQKCVLSQPIYVNAQ 484  
 QY 433 DTLSGTCIIILANKQSDISIAQVVDOTG-----SKSNLTLKNPFRYGTTPSP----- 479  
 DB 485 QEITGRHLIASQSYTINLTLSAKMGPANOGGIIOTSSCKLDLKEPYRMSQPAQVY 544  
 QY 480 TTSPSPS 488  
 DB 545 PTOEPNOS 553

RESULT 5  
 Q9S094 PRELIMINARY; PRT; 390 AA.  
 AC Q9S094:  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE ARGININE METHYLTRANSFERASE (PAM1).  
 GN T16L4.20 OR AT4G29510.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Rose M., Hempel S., Entian K.-D., Bancroft I., Mewes H.W.,  
 RA Mayer K.F.X., Lemcke K., Scheller C.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL079344; CAB45311.1; -  
 DR EMBL: AL161575; CAB79709.1; -  
 DR Interpro: IPR000051; -



KW    Transferase; Methyltransferase.  
SQ    SEQUENCE    390 AA; 43893 MW; 9BE43387CD7AE0B2 CRC64;

|                           |        |                    |            |             |
|---------------------------|--------|--------------------|------------|-------------|
| Query Match               | 15.7%  | Score 498.5;       | DB 10;     | Length 390; |
| Best Local Similarly      | 36.4%; | Pred. No. 2.5e-31; |            |             |
| Matches 125; Conservative | 55;    | Mismatches 142;    | Indels 21; | Gaps 8;     |

```

QY 114 TRPHFCSFYILKTCNGHTLESVFSETESSAVQYFQFYGLSQQNMMDYVTKTY 173
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 44 TPQESMF-----DAGESADIAEVTDDT--TSADYIFPQSYSHGCIHEMLKDVYTKTY 95

```

```

Qy 174 QRALQNHDFKDKIYLDVGGSGILSFFAQQAQARKIITAVEASTMAQHAEVLVKSNNLT 233
 | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 96 QNVYQNKFLIKDKIYLDVGAGTGILSLFCARAKGAHVAVECSQADMAKREIVKANGFS 155

```

Oy 234 DRIVEIPGKKEEVSLE-EQVDIIISEPBGYMLFENRMLESYLHAK-KYLKPSGNMPTTG 291  
| | : : ||| : | : ||| : | | : | : | : | : |  
Db 156 DVIIVLKGRKEEIELTPPKVDYIISFNMGYFLFEENMDLSVLYARDKMVLVEGGVLPDKA 215

```
Oy 292 DYNLAPFTDQOLYMEOFTKANFRYQPSBHGVDLSALRGAADVETFRQPVVDTFEDIRILMA 351
 :|| |:::||:|::|:|||||::
Db 216 SLHHTAIED----SEYKEDIIEFWNSYYGFDMSCIKKKAME-----PLYDVTVDQNOIVT 266
```

```
QY 352 KS-VKTYVFLLEAKEDDLHRIETIPFKHEMLHSLGVHGLAFWEDVAFIGSINTVMTSTAPT 410
 | : | : : | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 267 DSRLLKTMDSKMSSDA-STTAPFKLVAQRNDYIHALVAFYDVSTMCNKLGLGFSTGPK 325
```

```

Qy 411 EP LTHWYQYRCLFQSPLEFAKAGDTLSGICLLIANKRQSDIST 453
 ||| | : | :|::|| : || |
Db 326 SRATHMKQFVLYLDEDVLTICESETITTCMSVSPKKKNRPDIDI 368

```

RESULT 6  
081813

|    |                                                   |
|----|---------------------------------------------------|
| AC | 081813;                                           |
| DT | 01-NOV-1998 (TTEMBLrel. 08, Created)              |
| DT | 01-NOV-1998 (TTEMBLrel. 08, Last sequence update) |

DE ARGININE METHYLTRANSFERASE (FRAGMENT).  
GN PAM1.  
OS *Arabidopsis thaliana* (Mouse-ear cress).

0C Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
0C Brassicales; Brassicaceae; Arabidopsis.  
0X NCBI\_TaxID=3702;

RP SEQUENCE FROM N.A.  
RA Salchert K.;  
RL Thesis (1997), Institut fuer Zuechtungsforschung,

|    |                                   |
|----|-----------------------------------|
| RN | [2]                               |
| RP | SEQUENCE FROM N.A.                |
| RA | Salchert K., Breuer F., Koncz C.; |

RT methyltransferase from *Arabidopsis thaliana*,"  
RL submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ007582; CAA07570.1; -

|    |                                 |         |           |                         |
|----|---------------------------------|---------|-----------|-------------------------|
| SO | SEQUENCE                        | 376 AA: | 42077 MW: | 32499E5E38A3D492 CRC64: |
| FT | NON_TER                         | 376     | 376       |                         |
| KW | Transferase; Methyltransferase. |         |           |                         |

|                       |        |              |        |             |
|-----------------------|--------|--------------|--------|-------------|
| Query Match           | 15.7%; | Score 496.5; | DB 10; | Length 376; |
| Best Local Similarity | 36.4%; | Pred No 3    | 4e-31. |             |

114 TPRHFCFYNILKTCRHTLERSVFSEPTLESSAVQFQFYGLISQOQNMMDYVRGTGTY 1733  
 matches 123; conservative 34; mismatches 143; indels 21; gaps

```

Db 44 TPQDESMF-----DAGESADTAEVTDIT--TSADYFDSYSHFGIHEEMLKDVVRTKTY 95

```

```
QY 174 QRALQNHTFDKDKILVDVCGSGILSEFAAQAGARKIYAVEASTMAQHAEVLVSNNLT 233
 | | | | | | | | | | | | | | | | | | | | : | | | | :
Db 96 QNVTYQNKFLLKDKILVDVGAGTGILSLFCAXGAGAHHYVAVECSQMADMKAEIVKANGFS 155
```

QY 234 DRIVIPGKVEVSLP-EQVDIIISEPGYMLFNERMLESYLHAK-KYLKPSGMMFPTIG 291

Db 156 DVITVLKGGKEIEITLPKVDIIISENGYFLLENNMDSVLVADKWLVEGGVLPDKA 215

```

Qy 292 DVHAPFTDQLYMEDFTKANFRYQPSFHGVDSALRGAAVDEYFRQPVVDTFDIRLIMA 351
 :|| - : : : :|| : : : :||| : :
Db 216 SLHLTAIED----SEYKEDKIEFWNSYVGFDMSCIKKRAMME-----PLVDTVDQNDIVT 266

```

```

Qy 352 KS-KYTVNPLEAKEDDLHRIEIPFKFPHMLHSGLVHGLAFWEDVAFIGSIWTVLSTAPT 410
 | : | : : | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 267 DSRLLKTMDSKMSGDA-SFAPAFKLVQANDYIHAIYAYFDVSYTMCHRLGFSIGPK 322

```

```

0y 411 EPLHHYQVNCLEFQSPLEAKAGDTLSCGCLLIANKRQSDISI 453
 ||| : | :|:|:| : || |
Db 326 SRATHHWKQVLYLEEDVLTICEGETITGTMSVSPAKKNPRIDI 368

```

| RESULT | 7 |
|--------|---|
| Q9VGM7 |   |

|    |                                                   |
|----|---------------------------------------------------|
| AC | Q9YCW7;                                           |
| DT | 01-MAY-2000 (TREMBLrel, 13, Created)              |
| DT | 01-MAY-2000 (TREMBLrel, 13, Last sequence update) |

DE CG6554 PROTEIN.  
GN CG6554.  
OS *Drosophila melanogaster* (Fruit fly).

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Ephydriidae; Ephydriinae;  
OC NCBI\_TaxID=7227;

RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.  
RA Foster C., Gabrielian A.E., Gard N.S., Gelbart W.M., Glasser K.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003688; AAF54556.1;  
 DR FlyBase: FBgn0037834; CG6554.  
 DR InterPro: IPR000051;  
 SO SEQUENCE 376 AA; 42805 MW; 9C6IDA1905619D44 CRC64;

Query Match 15.1%; Score 478; DB 5; Length 376;  
 Best Local Similarity 33.5%; Pred. No. 9.9e-30;  
 Matches 122; Conservative 63; Mismatches 149; Indels 30; Gaps 8;

DB 114 TPHDGCFYNILK--TCRGHTLEKVSFSESSAVQYFOFYGYLSQOQNMMDYVTRG 171  
 19 TPNSNANNNVAKKLPAEGSTGD-NPNANADEMTSRDYFDSYAHFGIHEMLDEVRTV 77  
 QY 172 TYQRAILQNHTEKDYLDVGGCGSGLISFPAAQAGARKITAVEASTMAQHAENVLYKSN 231  
 78 TYRANAMHNNHLEFGKTVLDVGGCTGLISMPAAKAGAAQVIAVDCSNITIEFARQVYIDNN 137  
 DB 232 LTRRIVVPEKVEVSP---EVDIITSEPMGYMLFERNLESTLHAK-KYLKPSGNMF 287  
 138 LQDVIYVAKGKIELELPNGIEVDIITSEWNGCLFESMLDVLVARDKWLKDKGMF 197  
 QY 288 PFDGVHAFPTDQOLYMEQFTKANFRYQPSFHGVDSALNGAIVDEFROVPDTPRIR 347  
 198 PDROTKLITLIED-----RQYKDEKIMMDVYGFDMKCIKRVATE---PLVDVDPK 248  
 DB 348 ILMAKSVKYVYNLEAKEGDLHRIE-----IPKFEHMLSGLVHGLAFEDVAFIGSI 400  
 249 QVNSTSCM-----VEVDLYTVOKADLNFSSKFSCLIKHNDFQALVTFENIEFTKCH 301  
 QY 401 MTWVLSAPTEPLTHWYOVRCLFQSPLEAKGDTLSGTCLLIANKRSQVDSIVAQVOT 460  
 302 KRLEFSTSPDSTYTHMKQYFYLDDHMTAKNKEITGFQMKPNERNRNDLVIDINFK 361  
 DB 461 GSKS 464  
 362 GELS 365

RESULT 8  
 Q9MAT5 PRELIMINARY; PRT; 383 AA.  
 AC Q9MAT5;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE F13M7.14 PROTEIN.  
 GN F13M7.14.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_Taxid=3702;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Kwan A., Liu, S.,  
 RA Li J., Kremetskaia I., Luros J., Araujo R., Au M., Bredel V.,  
 RA Buehler E., Conway A., Dewar K., Feng J., Kim C., Kutz D., Li Y.,  
 RA Palm C., Shinn P., Sun H., Davis R., Ecker J., Federpfeil N.,  
 RA Theologis A.;  
 RT "The sequence of BAC F13M7 from Arabidopsis thaliana chromosome 1";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Theologis A.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Theologis A.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Theologis;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC004809; AAF40450.1;  
 DR InterPro: IPR000051;  
 SO SEQUENCE 383 AA; 43129 MW; 9EP6BDF7EE30B3F5 CRC64;

Query Match 15.1%; Score 477.5; DB 10; Length 383;  
 Best Local Similarity 33.6%; Pred. No. 1.1e-29;  
 Matches 125; Conservative 66; Mismatches 124; Indels 57; Gaps 12;

QY 143 ESSAVQYFOFYGYLSQOQNMMDYVTRGTYQRAILQNHTEKDYLDVGGCGSGLISFF 202  
 28 KEVDYAQYFCYYSFLYHOKMDLSDVRMDAYFNAYFQNHHEFGKTVLDVGGSGILAI 87  
 DB 203 AAQAGARKITAVEASTMAQHAENVLYKSNLIDRIYVTPKVEVSLPEQVDIITSEPMGY 262  
 88 SAQAGARVYAVEATKMDHARALVKNLIDHIVEITBSVEDISLPEKVDVITSEWNGY 147  
 QY 263 MLFERNLESTLHAK-KYLKPSGNMFPTIGVHLAP-----FT 299  
 148 FLIRSMFDVYSADRNKLPFGVWYPSHARWMLAPIKSNIDRRKNDPDGAAADWHNS 207  
 QY 300 DEQLYMEQFTKANFRYQPSFHGVDSALNGAIVDEFROVPV--DTFDIRILMAKSV 354  
 208 DE-----IKSYGVDMGVTLKPPAEDEQKYYIOTAMNDLNPOLIGTPTI 253  
 QY 355 KYTVNLEAKESDLH--RIETPFKFEHMLSGLVHGLAFEDVAFIGS---SIMTVWLS 407  
 254 VKEMCLTRASVEIEVRSNVTSVINMERTLC-GEQGFVDQFSGRKDPACQELTET 312  
 DB 408 APTPE-LTHWYOVRCLFQSPLEAKGDTLSGTCLLIANKRSQV--DISIVAVDQTSK 463  
 313 APSEQCHTHWGQVFIIMSPINVEEGDNLN-LGLMRSRKEHRLMEIELNCEIKRASGN 371  
 QY 464 SNLLDLKPNPF 475  
 372 PKE--SPKTYF 381

RESULT 9  
 Q9N8C5 PRELIMINARY; PRT; 345 AA.  
 AC Q9N8C5;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE POSSIBLE HNRNP ARGININE N-METHYLTRANSFERASE.  
 GN CHRL397.  
 OS Trypanosoma brucei.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 NCBI\_Taxid=5691;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TREU927;  
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,  
 RA Leonard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,  
 RA Gerrard C., Rajadream M.A., Barrell B.G.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL359782; CAB95620.1;  
 DR InterPro: IPR000051;  
 KW Transferase; Methyltransferase.  
 SO SEQUENCE 345 AA; 39206 MW; DECC06004F8AB776 CRC64;

Query Match 14.8%; Score 470; DB 5; Length 345;

Best Local Similarity 35.2%; Pred. No. 3.7e-29;  
Matches 124; Conservative 48; Mismatches 138; Indels 42; Gaps 8;

QY 132 TLENSVFSETESSAVOYFOFYGLSOOQNMODYRTGTQORAILQNHTEFKRVID 191  
DB 2 TVDANAASSTT--TTTDDYFDSYSHYGHMEMLKDCRHTTSRDAMWRNAYLFKDKRVID 59  
QY 192 VCGSGGIIISFPAAQAGARIVAVEASTMAQHAELVYKSNLDRIVYVIGKVEEVLPRQ 251  
DB 60 VCGCGGIIISFPAAQAGARIVAVEASTMAQHAELVYKSNLDRIVYVIGKVEEVLPRQ 119  
QY 252 VDIISSEPMYVLFNFRMLSEYLAHAKYL-KPSGNMFPTIGDVLHAPFDEOLYMEQFTR 310  
DB 120 VDIISSEPMYVLFNFRMLSEYLAHAKYL-KPSGNMFPTIGDVLHAPFDEOLYMEQFTR 176  
QY 311 ANFRYQPSFPGVDSALRGAADVDFRQPVDT-----FDIRILAKSVK 355  
DB 177 --FIEPMDVNOGIDFSEFKLS---FIEPLVDVERSOIVTNVAPLVSPDINTVKRADLS 230  
QY 356 YTVNF-LEAKEDDLRIEIPFEMHLSG---LVHGLAFWDVAFISIMYVWLSTAPF 410  
DB 221 FTSEFALRQASRGR-----NGNSIITYHALSVHDDPTAGHEVVIDTPY 280  
QY 411 EPLTHYQVRCLOFQSPDLPAKAGDTLSGCTLLANKRQSYDISIVAQDQGS 462  
DB 281 SPETHWROTLYLFPNLRAGERATFRMKCSFNALNGHDLISLHVDREGA 332

## RESULT 10

082210 PRELIMINARY; PRT: 366 AA.  
AC 082210;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
GN PUTATIVE ARGININE N-METHYLTRANSFERASE.  
CM F6622.30.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,  
RA Shen M., Ronning C.M., Fraser S.M., Somerville C.R., Venter J.C.,  
RA "Arabidopsis thaliana chromosome II BAC F6622 genomic sequence."  
RT Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AC005169; AAC62148.1; --  
DR InterPro: IPR000051; --  
KW Transferase; Methyltransferase.  
SQ SEQUENCE 366 AA; 41171 MW; 9C31A654BDC4C70 CRC64;

Query Match 14.8%; Score 467.5; DB 10; Length 366;  
Best Local Similarity 35.9%; Pred. No. 6.4e-29;  
Matches 112; Conservative 54; Mismatches 133; Indels 13; Gaps 6;

QY 145 SAAVQYFOFYGLSOOQNMODYRTGTQORAILQNHTEFKRVIDVYVIGKVEEVLPRQ 204  
DB 43 TSADYFDSYSHYGHMEMLKDCRHTTSRDAMWRNAYLFKDKRVIDVYVIGKVEEVLPRQ 102  
QY 205 QAGARIVAVEASTMAQHAELVYKSNLDRIVYVIGKVEEVLPRQ-EQVDIISSEPMYV 263  
DB 103 KAGAAHVAVESQAMDTAKETVKSNGFSDVITVLKGIIEILPVPKVDIISSEPMYV 162  
QY 264 LFNFRMLSEYLAHAK-KYLRPSGNMFPTIGDVLHAPFDEOLYMEQFTRANKRQSYDIS 322  
DB 163 LFNFRMLSEYLAHAK-KYLRPSGNMFPTIGDVLHAPFDEOLYMEQFTRANKRQSYDIS 217  
QY 323 DLSALRGAADVDFRQPVDTF-DIRILAKSVKITYVNFLEAKEDDLRIEIPFEMHLS 381

DB 218 DMSCKIRRAITE----PLVDTVDNQIVTDSKLLKTMIDISKMAAGDA-SFTAPKVLQOR 272  
QY 382 SGLVHGLAFWDVAFISIMYVWLSTAPFPLTHWYQVRCLOFQSPDLPAKAGDTLSGCTLL 441  
DB 273 NDHIALVAFVDFSFTRCHKRMKGSTGPKSRATVHMKTVLVLELDVLTICEGEITIGSMTI 332  
QY 442 IANKRQSYDIS 453  
DB 333 AQNKKRPRVDI 344

## RESULT 11

09VOX9 PRELIMINARY; PRT: 355 AA.  
AC 09VOX9;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE CG3675. PROTEIN.  
GN CG3675.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed-10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Abmayyan A., An H.-J., Andrews-Planck C., Baldwin D.,  
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Crawley S., Dahike C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
RA Talali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Pollard J., Puri V., Reese M.G.,  
RA Palazzolo M., Pittman G.S., Pan S., Saunders R.D.C., Scheefel F., Shen H.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheefel F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Szelecs R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
DR EMBL: AE003577; AAF51032.1; --  
DR FlyBase: FBgn0031592; CG3675.  
DR InterPro: IPR000051; --  
SQ SEQUENCE 355 AA; 40452 MW; 6AF252AE284F94A CRC64;

```

Query Match 14.3%; Score 453.5; DB 5; Length 355;
Best Local Similarity 33.8%; Pred. No. 7.9e-28;
Matches 117; Conservative 66; Mismatches 130; Indels 33; Gaps 10;

QY 131 HTLEBSVFSRTSESSAVOYFOFYGYLSQQOONMODVRYCTYORALQNTDKDKYVL 190
 ||| :
Db 16 HTANQITIKRRROEE--HFEKLGRIEHEMLKDSRIKAYKEAI-QHNEFFRHKVL 72

QY 191 DVGCGSGLISFFAQAQARKIYAEASTMAQHAELVKSNNLDRIVIPGKVEESLPE 250
 ||||| :
Db 73 DVGCGSGLISFFAQAQARKIYAEASTMAQHAELVKSNNLDRIVIPGKVEESLPE 132

QY 251 ---QVDIIISPEMGYMLFNERMLESTYLHAK-KYIKPSGNNFPTIGVHLAFT--TDEQLY 304
 ||||| :
Db 133 GIKKVDIIVCDWMSGSCFFSGNMLESILFARDKMLSATGHIYPTDPAQYLAIKGRDQDL- 191

QY 305 MEQFTKANFYQPSFHFVDLSALR---GAAVDEYFRQPVYDTRIDRIIMAKSVK--TVN 359
 ||||| :
Db 192 -----GFHWDHGFDSLAIKRCESKAVVEHV--TDDQMSRVCILSLDLYTER 239

QY 360 FLEAKEGDLHRIEIPKFKHMLHSGVLGHLAFWEDVAFISIMTWLSTAPETPLTHWYQV 419
 ||| :
Db 240 ROSAKSRSL-----PELXYSRNGWAGVLAIVDVGFSKSRQISFSTSPAPPTHNNQT 293

QY 420 RCLFQSPFLFAKAGDTLSGCLLIANKROSIDISIVAOVDQTSKSS 465
 ||| :
Db 294 VFLETPPLPVRAAGEICIKVLTMRPSEDSIDFTEFDIEVNFEDGRKSS 339

RESULT 12
Q9NVR8 PRELIMINARY; PRT; 316 AA.
AC 09NVR8;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE CDNA FLJ10359 FIS, CLONE NT2RP2002618, WEAKLY SIMILAR TO PROTEIN
DE ARGININE N-METHYLTRANSFERASE 2 (EC 2.1.1.-).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara K., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT "NEO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001421; BAA91681.1;
DR InterPro: IPR000051;
SQ SEQUENCE 316 AA; 35238 MW; E35E5567C2EAB0E CRC64;

```

```

Query Match 14.3%; Score 452.5; DB 4; Length 316;
Best Local Similarity 38.1%; Pred. No. 7.9e-28;
Matches 117; Conservative 39; Mismatches 108; Indels 43; Gaps 8;

QY 163 MMODVYRTGYORALQNTDKDKYLDVCGSGSILSFQAQAQARKIYAEASTMAQH 222
 ||| :
Db 1 MMDVYRTGYORALQNTDKDKYLDVCGSGSILSFQAQAQARKIYAEASTMAQH 60

QY 223 AEVLVKSNNLDRIVIPGKVEESLPEQVDIIISPEMGYMLFNERMLESTYLHAK-KYLK 281
 ||| :
Db 61 AREVYRNGEDREVYHVPGEYVELPEQVDIYSEMGYGLLHESMLSLVLAHARTKWL 120

QY 282 PSGNMFTIGDVHLAFTDQLYMEQFTKANFYQPSFH--GVDSLALGGAVIDEYFRPV 340
 ||| :
Db 121 EGGILLPASAELEFAPISDQMLEW-----RLGFSQVAKOHGVDVNSCLEG----- 165

QY 341 VDFEDRIILMAKSVKYVNFLEAKEG-----DLHRIEIP-----FKFHL 380

RESULT 13
Q9SNQ2 PRELIMINARY; PRT; 380 AA.
AC 09SNQ2;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE SIMILAR TO PROTEIN ARGININE N-METHYL TRANSFERASE 1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae;
OC Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GAJ) genomic DNA, chromosome 6, PAC
RT clone: p0535G04."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP000399; BAA83575.1;
DR InterPro: IPR000051;
DR InterPro: IPR001189;
DR ProDom: PD000475;
KW Transferase.
SQ SEQUENCE 380 AA; 42717 MW; B857E3153279CA0F CRC64;

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 20, 2001, 09:55:06 ; Search time 13.98 Seconds  
(without alignments)  
876.104 Million cell updates/sec

Title: US-09-464-377-2

Perfect score: 3168  
Sequence: 1 MAAAAATVGPAGSAGVAG.....PAISNAPSMTPTNTHYGS 608

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 2014635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description       |
|------------|--------|-------------|--------|-------|-------------------|
| 1          | 2290.5 | 72.3        | 447    | 3     | US-09-109-204-3   |
| 2          | 481.5  | 15.2        | 343    | 3     | US-09-109-204-32  |
| 3          | 127.5  | 4.0         | 276    | 4     | US-08-935-263-16  |
| 4          | 117    | 3.7         | 531    | 2     | US-08-923-536A-12 |
| 5          | 114    | 3.6         | 1261   | 4     | US-09-208-742-4   |
| 6          | 108    | 3.4         | 356    | 1     | US-07-959-941-2   |
| 7          | 108    | 3.4         | 356    | 1     | US-08-259-924-2   |
| 8          | 107.5  | 3.4         | 1289   | 1     | US-07-876-280-4   |
| 9          | 107.5  | 3.4         | 1289   | 1     | US-07-675-772-4   |
| 10         | 107.5  | 3.4         | 1289   | 1     | US-08-063-170-4   |
| 11         | 107.5  | 3.4         | 1289   | 1     | US-08-158-232-4   |
| 12         | 107.5  | 3.4         | 1289   | 1     | US-08-304-626-4   |
| 13         | 107.5  | 3.4         | 1289   | 1     | US-08-316-301A-4  |
| 14         | 107.5  | 3.4         | 1289   | 2     | US-08-611-928-4   |
| 15         | 107.5  | 3.4         | 1289   | 3     | US-09-173-891-4   |
| 16         | 107.5  | 3.4         | 1289   | 4     | US-09-076-137-4   |
| 17         | 107.5  | 3.4         | 1289   | 5     | PCT-US92-03624-4  |
| 18         | 107.5  | 3.4         | 1289   | 6     | 5281530-3         |
| 19         | 107.5  | 3.4         | 1289   | 6     | 5426049-4         |
| 20         | 103.5  | 3.3         | 696    | 3     | US-08-906-865-4   |
| 21         | 99     | 3.1         | 552    | 1     | US-07-662-223-2   |
| 22         | 98.5   | 3.1         | 3665   | 2     | US-08-222-617A-13 |
| 23         | 98.5   | 3.1         | 3712   | 2     | US-08-222-617A-4  |
| 24         | 98.5   | 3.1         | 3712   | 2     | US-08-222-617A-25 |
| 25         | 96.5   | 3.0         | 675    | 1     | US-08-386-495-10  |
| 26         | 96.5   | 3.0         | 675    | 5     | PCT-US96-02331-10 |
| 27         | 95     | 3.0         | 654    | 1     | US-08-083-590A-16 |

|    |      |     |      |   |                   |                   |
|----|------|-----|------|---|-------------------|-------------------|
| 28 | 95   | 3.0 | 654  | 2 | US-08-346-128-35  | Sequence 35, Appl |
| 29 | 95   | 3.0 | 654  | 3 | US-08-532-384-16  | Sequence 16, Appl |
| 30 | 95   | 3.0 | 2523 | 1 | US-08-185-432-18  | Sequence 18, Appl |
| 31 | 93.5 | 3.0 | 661  | 1 | US-08-525-742-2   | Sequence 2, Appl  |
| 32 | 92.5 | 2.9 | 666  | 1 | US-08-083-590A-17 | Sequence 17, Appl |
| 33 | 92.5 | 2.9 | 666  | 2 | US-08-346-128-36  | Sequence 36, Appl |
| 34 | 92.5 | 2.9 | 666  | 2 | US-08-532-384-17  | Sequence 3, Appl  |
| 35 | 92.5 | 2.9 | 1118 | 2 | US-08-724-354D-2  | Sequence 2, Appl  |
| 36 | 92.5 | 2.9 | 1118 | 2 | US-09-270-984A-2  | Sequence 8, Appl  |
| 37 | 92   | 2.9 | 1452 | 2 | US-08-449-644-8   | Sequence 8, Appl  |
| 38 | 92   | 2.9 | 1452 | 2 | US-08-087-244A-8  | Sequence 112, App |
| 39 | 91.5 | 2.9 | 436  | 3 | US-09-188-579-112 | Sequence 3, Appl  |
| 40 | 91.5 | 2.9 | 436  | 4 | US-09-315-444-112 | Sequence 3, Appl  |
| 41 | 91   | 2.9 | 240  | 1 | US-07-965-668A-3  | Sequence 3, Appl  |
| 42 | 91   | 2.9 | 240  | 2 | US-08-950-433-3   | Sequence 3, Appl  |
| 43 | 91   | 2.9 | 240  | 3 | US-09-186-287-3   | Sequence 25, Appl |
| 44 | 91   | 2.9 | 504  | 1 | US-08-457-274A-25 | Sequence 25, Appl |
| 45 | 91   | 2.9 | 504  | 5 | PCT-US95-05758-25 | Sequence 25, Appl |

## ALIGNMENTS

RESULT 1  
US-09-109-204-3  
Sequence 3, Application US/09109204  
Patent No. 6060250  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Gorgone, Gina A.  
APPLICANT: Corley, Neil C.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: HUMAN TRANSFERASES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/109,204  
FILING DATE: HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: -  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cerrone, Michael C  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0546 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-855-0572  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRATUT21  
CLONE: 2525071

US-09-109-204-3

Query Match 72.3%; Score 2290.5; DB 3; Length 447;  
 Best Local Similarity 98.7%; Pred. No. 4.9e-205;  
 Matches 441; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 163 MMODVYRTGYQRAIQNTDFKDKIVLDVGGSGILSFFAAOAGAKKIYAVASTMAOH 222  
 DB 1 MMODVYRTGYQRAIQNTDFKDKIVLDVGGSGILSFFAAOAGAKKIYAVASTMAOH 60  
 QY 223 AEVLVKSNNLTORIVYIPKKEVEEVSLEPOVDIISSEPMGMLENEMLESYLAHKKYLP 282  
 DB 61 AEVLVKSNNLTORIVYIPKKEVEEVSLEPOVDIISSEPMGMLENEMLESYLAHKKYLP 120  
 QY 283 SGNMFPPTIGDVLHAPFTDQLYMEQFTKANFRYQPSFGVDSLALGAAVDEYFROPVVD 342  
 DB 121 SGNMFPPTIGDVLHAPFTDQLYMEQFTKANFRYQPSFGVDSLALGAAVDEYFROPVVD 180  
 QY 343 TFDIRLMAKSKYKYTNFLEAKEGDLHRIEIPKKEHMLHSGLVHGLAFWEDVAFISIMT 402  
 DB 181 TFDIRLMAKSKYKYTNFLEAKEGDLHRIEIPKKEHMLHSGLVHGLAFWEDVAFISIMT 240  
 QY 403 VMLSTAPTEPLTHWYOVRCLFOSPLFAKAGDTLSTCLLIANKROSDISIVAOVDQTS 462  
 DB 241 VMLSTAPTEPLTHWYOVRCLFOSPLFAKAGDTLSTCLLIANKROSDISIVAOVDQTS 300  
 QY 463 KSNLLDLNPFPRYTGTPSPPGSHYTSPEENMNTSTYNLSSGVAVAGMPTAYDLS 522  
 DB 301 KSNLLDLNPFPRYTGTPSPPGSHYTSPEENMNTSTYNLSSGVAVAGMPTAYDLS 360  
 QY 523 SYIAGSSVGHNNLPLANTGIYNHSHRSGSJMSTGIYOGSSGAGG-GGGSSSAHAYN 561  
 DB 361 SYIAGSSVGHNNLPLANTGIYNHSHRSGSJMSTGIYOGSSGAGGSGGSTSAHAYN 420  
 QY 582 NOFTMGPAISMASPMISPTNTMHYCS 608  
 DB 421 SOFTMGPAISMASPMISPTNTMHYCS 447

RESULT 2

US-09-109-204-32  
 Sequence 32, Application US/09109204  
 Patent No. 6060250

GENERAL INFORMATION:

APPLICANT: Lal, Preeti  
 APPLICANT: Bandman, Olga  
 APPLICANT: Hillman, Jennifer L.  
 APPLICANT: Guegler, Karl J.  
 APPLICANT: Gorgone, Gina A.  
 APPLICANT: Corley, Neil C.  
 APPLICANT: Patterson, Chandra  
 TITLE OF INVENTION: HUMAN TRANSFERASES  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/109,204

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Cerrone, Michael C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PF-0546 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-855-0572

TELEX:

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 343 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: GI 1808648

US-09-109-204-32

Query Match 15.2%; Score 481.5; DB 3; Length 343;  
 Best Local Similarity 33.4%; Pred. No. 1.2e-36;  
 Matches 113; Conservative 65; Mismatches 129; Indels 31; Gaps 7;

QY 139 SEETESSAVQ-----YFQFYGLSQQMMODVYRTGYQRAIQNTDFKDKIVLDV 192  
 DB 8 AESSEKPNADMTSKDYTFPSYAFHGIHEMLKDEVRTLLYRNSMFHRLFKRVVLDV 67  
 QY 193 GCGSGILSFFAAOAGAKKIYAVEASTMAOHAELVKSNNLTDRIVIPGKVEEVSLEP-EQ 251  
 DB 68 GCGSGILCMFAAKAGAKKIYGIYVSSISDYAVKLVKANKLDHVVTIIKQKVEEVELPEK 127  
 QY 252 VDIISSEPMGMLENEMLESYLAH-KYKPSGNMFPPTIGDVLHAPFTDQLYMEQFTK 310  
 DB 128 VDIISSEPMGMLENEMLESYLAH-KYKPSGNMFPPTIGDVLHAPFTDQLYMEQFTK 182  
 QY 311 ANFRYQPSFGVDSLALGAAVDEYFROPVYDFDIRLMAKSKYKYTNFLEAKEGDLH 370  
 DB 183 YKIHMEVNYGFDMSCKDAI-----KEPLVDVYDPKQV-----TNCLIKVEVDIYT 231  
 QY 371 IEI-----PKFHMHLHSGLVHGLAFWEDVAFISIMTVMSTAPTEPLTHWYOVRCLF 423  
 DB 232 VKVEDLFTSPFLQVVRNDYVHALVAFNIEFRCRKRTGFSPEPYTHMKQTVAYM 291  
 QY 424 QSPLEAKAGDTLSTCLLIANKROSDISIVAOVDQTS 461  
 DB 292 EDYLTVTGTGEIFGTIGMRPNKNNRDLDTIDDFKG 329

RESULT 3

US-08-935-263-16  
 Sequence 16, Application US/08935263A  
 Patent No. 6117669

GENERAL INFORMATION:

APPLICANT: Furuichi, Yasuhiro  
 APPLICANT: Hoshino, Tatsuo  
 APPLICANT: Kimura, Hitoshi  
 APPLICANT: Kiyasu, Tatsuya  
 APPLICANT: Nagahashi, Yoshie  
 TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES  
 FILE REFERENCE: Biotin Genes  
 CURRENT APPLICATION NUMBER: US/08/935,263A  
 CURRENT FILING DATE: 1997-09-22  
 EARLIER APPLICATION NUMBER: EP 96115540.5  
 EARLIER FILING DATE: 1996-09-27  
 NUMBER OF SEQ ID NOS: 23  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 16

LENGTH: 276

TYPE: PRT

ORGANISM: Kurthia sp.

US-08-935-263-16

Query Match 4.0%; Score 127.5; DB 4; Length 276;

Best Local Similarity 20.0%; Pred. No. 0.00084; Matches 59; Conservative 52; Mismatches 87; Indels 97; Gaps 12;

QY 133 LERSVFSEETESSAVQYFOFYGLISQOQNMADYRTGTORAILQNHTEPKDKI-VLD 191

DB 2 IDKOLSKRSEHAKT-----YDAVAVQKNAKQVLD-----LTFQKNSKQRTINILE 49

QY 192 VGGGSLSEFAAAGARKIYAEASTMAQHAENVLYKSNMLDRIVIPGKVEEVSLEPQ 251

DB 50 ICGGTGLTRLVMTFPRNA-SITAVDLAPGMVEAKGITEEDRTFICADIEETLWEN 107

QY 252 VDIITSE-----PQGYMLFN-----ERMLSYLHAKKYLKPS 283

DB 108 YDLIISNATFQWLNLPGLTIEQLFTRLTPEGNLIFSTFGIKTFQELHMSYEHAKKQLS 167

QY 284 -----GMMFTIGDV-----HLAPFTD-----EQLYMEQF----- 308

DB 168 IDSSPGQLFALDELSCICEBAIPFSSAFLPTEIKTELELEYQTVREFFTSIKKIGAA 227

QY 309 --TRANRYOPSEFHGVLDLSALRGAAVDEYFRQPVVDFTDILMAKSVKYTVNFL 361

DB 228 NSKNENTCQPSF-----FRF-LINITYETKQDESGVKATYHCL 265

#### RESULT 4

US-08-923-536A-12

Sequence 12, Application US/08923536A

Patent No. 5965426

GENERAL INFORMATION:

APPLICANT: SAKAI, Yasuyoshi

APPLICANT: KATO, No. 5965426uo

APPLICANT: SHIBANO, Yui

TITLE OF INVENTION: PROTEIN DISULFIDE ISOMERASE GENE DERIVED

TITLE OF INVENTION: FROM STRAIN OF METHYLOTROPIC YEAST

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.

STREET: 1737 King Street, Suite 500

City: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22314-2756

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/923,536A

FILING DATE: 04-SEP-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 8-234287

FILING DATE: 04-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Meuth, Donna M.

REGISTRATION NUMBER: 36,607

REFERENCE/DOCKET NUMBER: 001560-317

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 531 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-923-536A-12

Best Local Similarity 21.2%; Pred. No. 0.023; Matches 81; Conservative 49; Mismatches 130; Indels 122; Gaps 17;

QY 63 AGPDAGIAL-----YSHEDVCYFK-----CSVSRETECSRVRGQ-----SFLITLG 104

DB 28 ASPDSAVVYKLTADSPESFKNRPVLAIEFFAPWCG-----HCKRLGPEFOVADKLVEMD 82

QY 105 CNSVLIOGATIPDFCSFPNI-----LKTGRHTLERSVSEETESSAVQYFOFYGLSQ 159

DB 83 IRLAQIDTEEKDLCSYGINKYPTLKVRGYENEPSDYGARTSDSIISYM----- 134

QY 160 QQNMADYVRCITYQRAILOHNTDFDKI-----VLVGGGS-----GLISFEAQA 206

DB 135 -----VKOSTPPVSTVDDLSIEDTKESNDPVFIQVLKSGKSVAGNSTFEILAN 186

QY 207 GARKIYAEASTMAQHAENVLYKSNMLDRIVIPGKVEEVSLEPEQDIITSEPMGYMLF- 265

DB 167 GLRDNYSFISTSTSEFSSKYLKG-----IKKSDTPSYILR 222

QY 266 -NEMLESTYLAHAKKYLKSGNFPITIGVHLAPFTDEQLYMEQFTKANFRYQPSFHGVDL 324

DB 223 PHEELSDASIV-----KFDEIDDTHLIEF-----LVNESKPLFEGEMDG 260

QY 325 SALRG-----AAVDYFRQPVVD-----TPDIRILMAKSVKYTVNF--LEAKEGDLH----- 369

DB 261 SSFQSYMEKMLPVAYFYFNEISEKDAVSDAISKLAKTHRGKVNFGDLSKYGGLHAKNIN 320

QY 370 -RIEIP-FKFMHLHSGLVHGLA 389

DB 321 MKEEPLRAIHDLATELKYGIS 342

#### RESULT 5

US-09-208-742-4

Sequence 4, Application US/09208742

Patent No. 6174679

GENERAL INFORMATION:

APPLICANT: Kaufmann, Joerg

TITLE OF INVENTION: C1F150/HrF1150 is Necessary for Cell

TITLE OF INVENTION: Cycle Progression

FILE REFERENCE: 1453.002

CURRENT APPLICATION NUMBER: US/09/208,742

CURRENT FILING DATE: 1998-12-10

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4

LENGTH: 1261

TYPE: PRT

ORGANISM: human

US-09-208-742-4

Query Match 3.6%; Score 114; DB 4; Length 1261;

Best Local Similarity 21.9%; Pred. No. 0.18; Matches 91; Conservative 64; Mismatches 167; Indels 94; Gaps 20;

QY 216 ASTMAQHAENVLYKSNMLDRIVIPGKVEEVS--LPEQVDIITSEPMGYMLFERNMLESY 273

DB 483 SATFFKKTIEKLAR-DILIDPIRYOGDIEANEDVQYIIEILHSGPSKMMWTLRLRIVE-- 539

QY 274 LHAKKYLPKSGMFPPTIGVHLAPFTDEQLYMEQFTKANFRYQPSFHGVDLSALRGAAVD 333

DB 540 -----FTSSGSLVLF--FVTKRANAEELAN-NIKOE-----GNNLGLLHG-DMD 578

QY 334 EYFRQPVVDTF--DIRILMA-----KSVKYTVNPLEKEGDLHRIEIPFKFHM 360

DB 579 QSERKKVVISDEKFKDIPVLAIVADVAARGLDIPSIKTVINVDVARDIDTHRIIGRTGRAG 638

QY 381 HSGLVHGLAFMFDVAFIGSIM--TWLSTAPTEPLTHWYQVRCLOFQSPPLFAKAGDTLSGT 438

DB 639 EKGVAVTLLTPKDSNFAGLVYRNLBGANOHVSKELLDLAMQNMWFKRSR-F-KGK 634

QY 439 CLIIA-----NKRQSTDISIVAQVDGTGSKSSNLDLKNPF-F 475

Query Match 3.7%; Score 117; DB 2; Length 531;



Db 695 KNTGGGGLYERERPOLSENNDRGNNTASNE-----AKPCTGAMGRITMKAPFOS 750  
OY 476 RYGTTPPPSHSTYSPSENNMTNGSTYNNLSSGVAAGMPTAYDISVYAGGSVGHN- 534  
Db 751 QYK-----SHFVAASLSNOKAG-----SSAAGASGWTSGNSINSVPTNSAQCHNS 796  
OY 535 --NLPLANTGIVNTHSRMSGMSIGIVGSSGAGGSSSAHYAVNNQFTMG 588  
Db 797 PDSPTSAKGIPIRGNT--GNISGAPVTPYPSAGAGVNNNTASGN--NSRSTGG 847

## RESULT 6

US-07-959-941-2  
; Sequence 2, Application US/07959941  
; Patent No. 5364781  
; GENERAL INFORMATION:  
; APPLICANT: HUTCHINSON, Charles R.  
; APPLICANT: MADDURI, Krishna M.  
; APPLICANT: TORTI, Francesca  
; APPLICANT: COLOMBO, Anna L.  
; TITLE OF INVENTION: PROCESS FOR PREPARING DAUNORUBICIN  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikaïdo, Marmelstein, Murray & Oram  
; STREET: 655 Fifteenth Street N.W. Suite 330.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/959,941  
; FILING DATE: 19921009  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 793,873  
; FILING DATE: 18-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Chiu, Monica F.  
; REGISTRATION NUMBER: P-36,105  
; REFERENCE/DOCKET NUMBER: 1615-1816CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)638-5000  
; TELEFAX: (202)638-4810  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 356 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-959-941-2

Query Match 3.4%; Score 108; DB 1; Length 356;  
Best Local Similarity 27.6%; Pred. No. 0.083;  
Matches 53; Conservative 26; Mismatches 67; Indels 46; Gaps 10;

OY 189 VLDVGGGSGILSFPAAGARKIYAVEASTM-----AQHAEVLVKSNNLTDRIIVIPGKV 243  
Db 183 VLDVGGGKG---GFAA-AIARRAPHSATVLEMAGYDTARSYLKDEGLSDRDVVEGDF 238  
OY 244 EEVSLPEVDYDIISPEMGYLFN-----ERMLESYLAHKYLLKPSGMMPTIGDVILAP 297  
Db 239 FE-PLPRKADAIL--LSFVLLNMPDHDVAVILTRCAEA---LEPGRIIL-----IHERD 286  
OY 298 FTDEQLYMEQFTKANFRYQSPFHGV-----DLSALRGAADVDEYFRQPVVDFTDIRIL 349  
Db 287 DLHENSFNEQFTELDRLMVLVFLGALTRREKWDGLAASAGLVVE-----VRQL 335

OY 350 MAKSVKTYVNF 361  
Db 336 PSPITPYDLSL 347

## RESULT 7

US-08-259-924-2  
; Sequence 2, Application US/08259924  
; Patent No. 5363064  
; GENERAL INFORMATION:  
; APPLICANT: HUTCHINSON, Charles R.  
; APPLICANT: MADDURI, Krishna M.  
; APPLICANT: TORTI, Francesca  
; APPLICANT: COLOMBO, Anna L.  
; TITLE OF INVENTION: PROCESS FOR PREPARING DAUNORUBICIN  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikaïdo, Marmelstein, Murray & Oram  
; STREET: 655 Fifteenth Street N.W. Suite 330  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/259,924  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 793,873  
; FILING DATE: 18-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/959,941  
; FILING DATE: 09-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kiltz, Monica C.  
; REGISTRATION NUMBER: 36,105  
; REFERENCE/DOCKET NUMBER: 1615-4003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)638-5000  
; TELEFAX: (202)638-4810  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 356 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-259-924-2

Query Match 3.4%; Score 108; DB 1; Length 356;  
Best Local Similarity 27.6%; Pred. No. 0.083;  
Matches 53; Conservative 26; Mismatches 67; Indels 46; Gaps 10;

OY 189 VLDVGGGSGILSFPAAGARKIYAVEASTM-----AQHAEVLVKSNNLTDRIIVIPGKV 243  
Db 183 VLDVGGGKG---GFAA-AIARRAPHSATVLEMAGYDTARSYLKDEGLSDRDVVEGDF 238  
OY 244 EEVSLPEVDYDIISPEMGYLFN-----ERMLESYLAHKYLLKPSGMMPTIGDVILAP 297  
Db 239 FE-PLPRKADAIL--LSFVLLNMPDHDVAVILTRCAEA---LEPGRIIL-----IHERD 286  
OY 298 FTDEQLYMEQFTKANFRYQSPFHGV-----DLSALRGAADVDEYFRQPVVDFTDIRIL 349  
Db 287 DLHENSFNEQFTELDRLMVLVFLGALTRREKWDGLAASAGLVVE-----VRQL 335  
OY 350 MAKSVKTYVNF 361



Db 656 NHHLTNGSSDVFLEDRIFEVPLESNTVTYIFNNSTYTGSAANLIPALIPALMSTSSDKALTG 715  
QY 509 GVAVAG-----MPTAYDLSSVIAGSSVGHNNLIPLANTGIVNHTSRMCSI 555  
Db 716 SMSITGRTTPNSDALLRFKFTNYDTQTIPIPGSGKDFNTLEIOD-----I 762  
QY 556 MSTGIYOGSSGAOGGSSSAHYAVNNOFTMGP 589  
Db 763 VSIDIFVG-SGLHSGDSIKLDFTNNSGSGGSP 795  
RESULT 10  
US-08-063-170-4  
Sequence 4, Application US/08063170  
Patent No. 5350576  
GENERAL INFORMATION:  
APPLICANT: Kim, Leo  
APPLICANT: Schwab, George E.  
TITLE OF INVENTION: Compositions and Methods for Inducing an Immune  
Response for Protection Against Endoparasites and Exoparasites  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DAVID R. SALIWANCHIK  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/063,170  
FILING DATE: 19930517  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,141  
FILING DATE: 30-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/759,248  
FILING DATE: 13-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: SALIWANCHIK, DAVID R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: M/S 103.C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1289 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Bacillus thuringiensis  
STRAIN: PS17  
INDIVIDUAL ISOLATE: PS17b  
US-08-063-170-4  
Query Match 3.4%; Score 107.5; DB 1; Length 1289;  
Best Local Similarity 21.9%; Pred. No. 0.76; Mismatches 97; Indels 81; Gaps 11;  
Matches 60; Conservative 36; Mismatches 97; Indels 81; Gaps 11;  
QY 383 GIVGLAFWFDVA-FISIMTVMSTLST-----PTE-----PLTH-----WYVRC----- 421  
Db 536 GNVSTMGPFPEKASYGTVKEMNGANAMKLSFGOSIGITITNTVTSGEYIRCRYASND 595

QY 422 -----LFOSPLEAKAGDTLSG-----TCLLANKROSYDISIAO-- 456  
Db 596 NTNFEFNDTGAGNPIFOQINFASVDDNNTGVGANGVYVKSATTDNSSFVKIPAKTI 655  
QY 457 -VDOTGSKSNLIDKQNFERYTGTTPSPPGSHYTPSEN-----MMNTGSTYNNLS 508  
Db 656 NHHLTNGSSDVFLEDRIFEVPLESNTVTYIFNNSTYTGSAANLIPALIPALMSTSSDKALTG 715  
QY 509 GVAVAG-----MPTAYDLSSVIAGSSVGHNNLIPLANTGIVNHTSRMCSI 555  
Db 716 SMSITGRTTPNSDALLRFKFTNYDTQTIPIPGSGKDFNTLEIOD-----I 762  
QY 556 MSTGIYOGSSGAOGGSSSAHYAVNNOFTMGP 589  
Db 763 VSIDIFVG-SGLHSGDSIKLDFTNNSGSGGSP 795  
RESULT 11  
US-08-158-232-4  
Sequence 4, Application US/08158232  
Patent No. 5596071  
GENERAL INFORMATION:  
APPLICANT: Payne, Jewel  
APPLICANT: Kennedy, W. Keith  
APPLICANT: Randall, John Brooks  
APPLICANT: Meier, Henry  
APPLICANT: Uick, Heidi Jane  
APPLICANT: Fonceerrada, Luis  
APPLICANT: Schnepf, H. Ernest  
APPLICANT: Schwab, George E.  
TITLE OF INVENTION: No. 5596071el Bacillus thuringiensis Toxins Active  
Against Hymenopteran Pests  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David R. Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/158,232  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/887,980  
FILING DATE: 22-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/797,645  
FILING DATE: 25-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/703,977  
FILING DATE: 22-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Saliwanchik, David R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: M/SC104.C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1289 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

```

MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: BACILLUS THURINGIENSIS
STRAIN: PS17
INDIVIDUAL ISOLATE: PS17b
IMMEDIATE SOURCE:
CLONE: E. coli NM522(pMYC1628) NRRL B-18652
US-08-158-232-4

```

```

Query Match 3.4%; Score 107.5; DB 1; Length 1289;
Best Local Similarity 21.9%; Pred. No. 0.76;
Matches 60; Conservative 36; Mismatches 97; Indels 81; Gaps 11;

```

```

QY 383 GLVHGLAFMFDA-FIGSIMTWLSTA-----PTE-----PLTH-----WYQVRC----- 421
DB 536 GNVSTMGPFPEKASVGGVYKEMLNGANAMKLSFGOSIGIPITNTVTSGETQIRCRVANSND 595
QY 422 -----LFQSPLEFAKAGDTLSG-----TCLLIANKROSVDISIVAQ-- 456
DB 596 NTNVEFNVDTGGANPIFOQINPASTVDNNTGVGANGVYVKSIAITDNTSFTYKIPAKTI 655
QY 457 -VDQTGSSNLLDKNPFRYGTTPSPPGSHYTSSEN-----MMNTGSTYVLS 508
DB 656 NVHLTNGSSDVFLDRIFEVPLIESNTVTIFNNSYTTGSANLIPALPIMSTSSDKALTG 715
QY 509 GVAVAG-----MPTAYDLSVYAGSSVGHNNLIPLANTGIYNHHSRSGSI 555
DB 716 SMSITGRTPNSDALLRFFKTYNDQTPIPPSGKDFNTLTLEID-----I 762
QY 556 MSTGIQSSGAGGSSSAHYAVNNOFTMGCP 589
DB 763 VSIDIFVG-SGLHSGDSIKLDFTNNSGSGGSP 795

```

```

RESULT 12
US-08-304-626-4
Sequence 4, Application US/08304626
Patent No. 5616495
GENERAL INFORMATION:
APPLICANT: Payne, Jewel M.
APPLICANT: Kennedy, M. Keith
APPLICANT: Randall, John Brooks
APPLICANT: Meier, Henry
APPLICANT: Uick, Heidi Jane
APPLICANT: Foncerrada, Luis
APPLICANT: Schnepf, Harry E.
APPLICANT: Schwab, George E.
TITLE OF INVENTION: No. 5616495el Bacillus thuringiensis Isolates
TITLE OF INVENTION: Active Against Hymenopteran Pests and Genes Encoding
TITLE OF INVENTION: Hymenopteran Active Toxins
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/304,626
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/887,980
FILING DATE:

```

```

ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: M/SCJ 104
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: BACILLUS THURINGIENSIS
STRAIN: PS17
INDIVIDUAL ISOLATE: PS17b
IMMEDIATE SOURCE:
CLONE: E. coli NM522(pMYC1628) NRRL B-18652
US-08-304-626-4

```

```

Query Match 3.4%; Score 107.5; DB 1; Length 1289;
Best Local Similarity 21.9%; Pred. No. 0.76;
Matches 60; Conservative 36; Mismatches 97; Indels 81; Gaps 11;

```

```

QY 383 GLVHGLAFMFDA-FIGSIMTWLSTA-----PTE-----PLTH-----WYQVRC----- 421
DB 536 GNVSTMGPFPEKASVGGVYKEMLNGANAMKLSFGOSIGIPITNTVTSGETQIRCRVANSND 595
QY 422 -----LFQSPLEFAKAGDTLSG-----TCLLIANKROSVDISIVAQ-- 456
DB 596 NTNVEFNVDTGGANPIFOQINPASTVDNNTGVGANGVYVKSIAITDNTSFTYKIPAKTI 655
QY 457 -VDQTGSSNLLDKNPFRYGTTPSPPGSHYTSSEN-----MMNTGSTYVLS 508
DB 656 NVHLTNGSSDVFLDRIFEVPLIESNTVTIFNNSYTTGSANLIPALPIMSTSSDKALTG 715
QY 509 GVAVAG-----MPTAYDLSVYAGSSVGHNNLIPLANTGIYNHHSRSGSI 555
DB 716 SMSITGRTPNSDALLRFFKTYNDQTPIPPSGKDFNTLTLEID-----I 762
QY 556 MSTGIQSSGAGGSSSAHYAVNNOFTMGCP 589
DB 763 VSIDIFVG-SGLHSGDSIKLDFTNNSGSGGSP 795

```

```

RESULT 13
US-08-316-301A-4
Sequence 4, Application US/08316301A
Patent No. 5753452
GENERAL INFORMATION:
APPLICANT: Schnepf, Harry E.
APPLICANT: Schwab, George E.
APPLICANT: Payne, Jewel M.
APPLICANT: Narva, Kenneth E.
APPLICANT: Foncerrada, Luis
APPLICANT: Saliwanchik & Saliwanchik
TITLE OF INVENTION: No. 5753492el Nematode-Active Toxins and Genes
TITLE OF INVENTION: Which Code Therefor
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/316, 301A  
FILING DATE: 30-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/871, 510  
FILING DATE: 23-APR-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/693, 018  
FILING DATE: 03-MAY-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/565, 544  
FILING DATE: 10-AUG-1990  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/084, 653  
FILING DATE: 12-AUG-1987  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/830, 050  
FILING DATE: 31-JAN-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lloyd, Jeff  
REGISTRATION NUMBER: 35, 589  
REFERENCE/DOCKET NUMBER: MA20CCCD1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1289 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: BACILLUS THURINGIENSIS  
INDIVIDUAL ISOLATE: PS17  
IMMEDIATE SOURCE:  
CLONE: E. coli NM522(PMYC 1628) NRRL B-18652  
US-08-316-301A-4

Query Match 3.4%; Score 107.5; DB 1; Length 1289;  
Best Local Similarity 21.9%; Pred. No. 0.76;  
Matches 60; Conservative 36; Mismatches 97; Indels 81; Gaps 11;

383 GLVHGLAFWFDVA-FIGSITWVLMSTF-----PTE-----PLTH-----WYOVRC----- 421  
DB 536 GNVSTMGFPPEKASYGTIVKKEWNGANMAKLSFGOSIGDITVNTSGEIVIRRIYASND 595  
422 -----LRFQSLFAKAGDTLSG-----TCLLANKROSYDISIVA-- 456  
DB 596 NTWVFVNDVGANPIRQINFSTVNNNGVYGANGVYVYKSIATIDNSTYVIRPKTI 655  
457 -VDOTGSKSSNLLDLKAPFRYGTTPSPPGSHYTSPESEN-----MNWTGSTYNLSS 508  
DB 656 NVHLTNGSSDVFDIRIEFVPILESNTVTTFNNSTYTGSAHLIPALAPLWSTSSDKALTG 715  
509 GVAVAG-----WPTAYDLSVYAGSSVGHNNLIPLANGIVVHHSRMGSI 555  
DB 716 SMSITGTTNSDALLRFKFTNTDTOTIPIGSGKDFNTLETOD-----I 762  
556 MSTGIYVGGSSGAOGGSSSAHYAVNNOFTWGP 589  
DB 763 VSDIDIFVG-SGLHSGDSIKLIDFTNNNSGSGSP 795

RESULT 14  
US-08-611-928-4  
Sequence 4, Application US/08611928  
Patent No. 5824792  
GENERAL INFORMATION:  
APPLICANT: Payne, Jewel  
APPLICANT: Kennedy, M. Keith  
APPLICANT: Randall, John Brooks  
APPLICANT: Meier, Henry  
APPLICANT: Vick, Heidi Jane  
APPLICANT: Foncerra, Luis  
APPLICANT: Schepf, H. Ernest  
APPLICANT: Schwab, George E.  
APPLICANT: Fu, Jenny  
TITLE OF INVENTION: No. 5824792el Bacillus thuringiensis Toxins Active  
TITLE OF INVENTION: Against Hymenopteran Pests  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David R. Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/611, 928  
FILING DATE: 06-MAR-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/158, 232  
FILING DATE: 24-NOV-1993  
APPLICATION NUMBER: US 07/887, 980  
FILING DATE: 22-MAY-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/797, 645  
FILING DATE: 25-NOV-1991  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/703, 977  
FILING DATE: 22-MAY-1991  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Saliwanchik, David R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: M/SCJ104.C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1289 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: BACILLUS THURINGIENSIS  
INDIVIDUAL ISOLATE: PS17b  
IMMEDIATE SOURCE:  
CLONE: E. coli NM522(PMYC1628) NRRL B-18652  
US-08-611-928-4



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2001, 16:46:39 ; Search time 2266.15 Seconds  
(without alignments)  
13031.227 Million cell updates/sec

Title: US-09-464-377-1

Perfect score: 3124  
Sequence: 1 agggggccttgagccggacc.....taaaagtgttccttgraa 3124

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1022815 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est121:\*  
191: gb\_est122:\*  
192: gb\_est123:\*  
193: gb\_est124:\*  
194: gb\_est125:\*  
195: gb\_est126:\*  
196: gb\_est127:\*  
197: gb\_est128:\*  
198: gb\_est129:\*  
199: gb\_est130:\*  
200: gb\_est131:\*  
201: gb\_est132:\*  
202: gb\_est133:\*  
203: gb\_est134:\*  
204: gb\_est135:\*  
205: gb\_est136:\*  
206: gb\_est137:\*  
207: gb\_est138:\*  
208: gb\_est139:\*  
209: gb\_est140:\*  
210: gb\_est141:\*  
211: gb\_est142:\*  
212: gb\_est143:\*  
213: gb\_est144:\*  
214: gb\_est145:\*  
215: gb\_est146:\*  
216: gb\_est147:\*  
217: gb\_est148:\*  
218: gb\_est149:\*  
219: gb\_est150:\*  
220: gb\_est151:\*  
221: gb\_est152:\*  
222: gb\_est153:\*  
223: gb\_est154:\*  
224: gb\_est155:\*  
225: gb\_est156:\*  
226: gb\_est157:\*  
227: gb\_est158:\*  
228: gb\_est159:\*  
229: gb\_est160:\*  
230: gb\_est161:\*  
231: gb\_est162:\*  
232: gb\_est163:\*  
233: gb\_est164:\*  
234: gb\_est165:\*  
235: gb\_est166:\*  
236: gb\_est167:\*  
237: gb\_est168:\*  
238: gb\_est169:\*  
239: gb\_est170:\*  
240: gb\_est171:\*  
241: gb\_est172:\*  
242: gb\_est173:\*  
243: gb\_est174:\*  
244: gb\_est175:\*  
245: gb\_est176:\*  
246: gb\_est177:\*  
247: gb\_est178:\*  
248: gb\_est179:\*  
249: gb\_est180:\*  
250: gb\_est181:\*  
251: gb\_est182:\*  
252: gb\_est183:\*  
253: gb\_est184:\*  
254: gb\_est185:\*  
255: gb\_est186:\*  
256: gb\_est187:\*  
257: gb\_est188:\*  
258: gb\_est189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.



## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 744.6 | 23.8        | 903    | 154   | BC468484 602510458 |
| 2          | 741.8 | 23.7        | 1883   | 152   | AK004959 Mus muscu |
| 3          | 727.6 | 23.3        | 1005   | 106   | AL530727 AL530727  |
| 4          | 695.4 | 22.8        | 943    | 106   | AL527554 AL527554  |
| 5          | 619.2 | 19.3        | 663    | 150   | BF580900 BF580900  |
| 6          | 610.6 | 19.5        | 668    | 151   | BF608742 BF608742  |
| 7          | 588.2 | 18.8        | 664    | 155   | BE292679 BE292679  |
| 8          | 577   | 18.5        | 577    | 6     | AA396116 AA396116  |
| 9          | 556.4 | 17.8        | 647    | 153   | BC386901 BC386901  |
| 10         | 555.2 | 17.8        | 919    | 153   | BC420290 BC420290  |
| 11         | 553.2 | 17.7        | 578    | 173   | BC063675 BC063675  |
| 12         | 552.8 | 17.7        | 568    | 3     | AA215095 AA215095  |
| 13         | 547.8 | 17.5        | 584    | 102   | AI839646 AI839646  |
| 14         | 533.4 | 17.4        | 713    | 147   | BF321245 BF321245  |
| 15         | 539.6 | 17.3        | 554    | 147   | AW538664 AW538664  |
| 16         | 539.4 | 17.3        | 607    | 147   | BF313667 BF313667  |
| 17         | 539.2 | 17.3        | 929    | 169   | BF794191 BF794191  |
| 18         | 538.6 | 17.2        | 591    | 104   | AI987990 AI987990  |
| 19         | 535.2 | 17.1        | 565    | 14    | AA980545 AA980545  |
| 20         | 530.8 | 17.0        | 629    | 166   | BE376609 BE376609  |
| 21         | 523.4 | 16.8        | 545    | 166   | BE311185 BE311185  |
| 22         | 522.6 | 16.7        | 739    | 165   | BE259059 BE259059  |
| 23         | 520.8 | 16.7        | 1027   | 172   | BF981587 BF981587  |
| 24         | 519.8 | 16.6        | 563    | 168   | BF722493 BF722493  |
| 25         | 517.8 | 16.6        | 555    | 23    | AI647525 AI647525  |
| 26         | 513.6 | 16.4        | 528    | 23    | AI642135 AI642135  |
| 27         | 511.6 | 16.4        | 611    | 165   | BE264306 BE264306  |
| 28         | 508.4 | 16.3        | 519    | 173   | BG093973 BG093973  |
| 29         | 503.4 | 16.1        | 778    | 137   | BE572052 BE572052  |
| 30         | 502.8 | 16.1        | 542    | 173   | BG077060 BG077060  |
| 31         | 501.2 | 16.0        | 956    | 168   | BF685687 BF685687  |
| 32         | 499.6 | 16.0        | 509    | 4     | AA272279 AA272279  |
| 33         | 498.6 | 16.0        | 529    | 149   | BF452047 BF452047  |
| 34         | 491   | 15.7        | 563    | 165   | BE253776 BE253776  |
| 35         | 477.4 | 15.3        | 489    | 145   | BF148817 BF148817  |
| 36         | 476.8 | 15.3        | 1109   | 140   | BE796260 BE796260  |
| 37         | 473.8 | 15.2        | 477    | 131   | W91119 W91119      |
| 38         | 473.8 | 15.2        | 489    | 166   | BE308812 BE308812  |
| 39         | 471.6 | 15.1        | 478    | 190   | W64905 W64905      |
| 40         | 460.4 | 14.7        | 739    | 166   | BE369209 BE369209  |
| 41         | 459.4 | 14.7        | 483    | 3     | AA162220 AA162220  |
| 42         | 457.6 | 14.6        | 540    | 119   | AW656460 AW656460  |
| 43         | 450.2 | 14.4        | 519    | 9     | AA596765 AA596765  |
| 44         | 448.8 | 14.4        | 517    | 120   | AW732639 AW732639  |
| 45         | 447   | 14.3        | 672    | 174   | BG160425 BG160425  |

## ALIGNMENTS

| RESULT     | 1                                                              | EST         | 21-MAR-2001     |
|------------|----------------------------------------------------------------|-------------|-----------------|
| LOCUS      | BC468484                                                       | 903 bp      | IMAGE:464446 5' |
| DEFINITION | 602510458F1 NIH_MGC_15 Homo sapiens CDNA clone                 |             |                 |
| ACCESSION  | BC468484                                                       | mRNA        |                 |
| VERSION    | BC468484.1                                                     | GI:13400754 |                 |
| KEYWORDS   | EST.                                                           |             |                 |
| SOURCE     | human                                                          |             |                 |
| ORGANISM   | Homo sapiens                                                   |             |                 |
| REFERENCE  | 1 (bases 1 to 903)                                             |             |                 |
| AUTHORS    | NIH-MGC http://mgc.nci.nih.gov/                                |             |                 |
| TITLE      | National Institutes of Health, Mammalian Gene Collection (MGC) |             |                 |
| JOURNAL    | Unpublished (1999)                                             |             |                 |
| COMMENT    | Contact: Robert Strausberg, Ph.D.                              |             |                 |

Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: NIH Intramural Sequencing Center  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 plate: LNCM1416 row: d column: 15  
 High quality sequence stop: 887.  
 Location/Qualifiers  
 1. 903  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:464446"  
 /clone\_lib="NIH\_MGC\_15"  
 /tissue\_type="adenoecarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: colon; Vector: pORF7; Site:1: XhoI; Site:2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGCG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 213 a 258 c 244 g 188 t  
 ORIGIN

Query Match 23.8%; Score 744.6; DB 154; Length 903;  
 Best Local Similarity 90.9%; Pred. No. 4.5e-190;  
 Matches 815; Conservative 0; Mismatches 79; Indels 3; Gaps 2;

499 cagcagcagacactgtatgcagagactatgtgcagacagacccaccagctgcacctg 558  
 |||||||  
 2 CACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 61  
 |||||||  
 559 cagaacacacacgagactcagaacagacagatcgttcagatgtggcctgtggatc 618  
 |||||||  
 62 CAAACACACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 121  
 |||||||  
 619 ctgtcattttttgtctgcacagcagcagcagcagcagcagcagcagcagcagcagc 678  
 |||||||  
 122 CTGTCTTTTTCCTGCGCCGACGACGACGACGACGACGACGACGACGACGACGACG 181  
 |||||||  
 679 atgtgtcagcagatgcagagctcgtgtgaagatgaacatctgcagacagcagctgtgc 738  
 |||||||  
 182 ATGGCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 241  
 |||||||  
 739 atccctggcgaagaatgaagagagctcattgttcgcgaagcgaatcattcattcagcag 798  
 |||||||  
 242 ATCCCGGCGGAGGTGAGGAGGTGACCTCCGAGAGGTGAGGACATCATCATCTCGAG 301  
 |||||||  
 799 cccatgtgtacatctcttcaatgaacgaatgtcagagatctacatctgcacaaag 858  
 |||||||  
 302 CCCATGGGCTACATGCTCTTCAACGAGCGGATCTGGAGGACCTCCGACGCAAG 361  
 |||||||  
 859 taccctgaagcctagtgaacatgttcccaacattgtgtatgttccacctgcagaccttc 918  
 |||||||  
 362 TACCTGAACCCGCGGAAACATGTTCCATCATGTTGATGACCTCCACCTTGCACCTTC 421  
 |||||||  
 919 actgatgaacagctctacatgtgagagatcaccagaagcagcttcggttcacagcttc 978  
 |||||||  
 422 ACGGATGAACAGCTCTACATGAGGAGGTGACGAGGACCAAGGACCACTTGTGTACCGACCACT 481  
 |||||||  
 979 ttccatgaagtggaacctgtcgcgacctcagaagtgccgctgtggatgagtaacttcgcgcaa 1038  
 |||||||  
 482 TTCATGAGGAGTGGACCTGTGCGCCGCGAGGTCGCGGCGGTGATGATTTTCGCGCAG 541  
 |||||||  
 1039 cctgtgtgtgaacatttgcacatcgcgagctcctgtatgtggccaacatctgtcaagttcacagt 1098  
 |||||||  
 542 CCTGTGTGTGACACATTTGACATCCGAGATCCTATGCGCCAAAGTCTGTCAAGTACAGCGTG 601  
 |||||||



| BASE COUNT | ORIGIN   |
|------------|----------|
| 174 a      | 320 c    |
| 320 c      | 347 g    |
| 160 t      | 4 others |

|                           |        |                     |           |              |
|---------------------------|--------|---------------------|-----------|--------------|
| Query Match               | 23.3%; | Score 727.6;        | DB 106;   | Length 1005; |
| Best Local Similarity     | 90.0%; | Pred. No. 1.8e-185; |           |              |
| Matches 788; Conservative | 3;     | Mismatches 82;      | Indels 3; | Gaps 1;      |

[illegible]

| REFERENCE          | ORGANISM                                                          | SOURCE | ALIMWORDS |
|--------------------|-------------------------------------------------------------------|--------|-----------|
| 1 (bases 1 to 943) | Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.       | human. | 251.      |
|                    | Homo sapiens                                                      |        |           |
|                    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |        |           |
|                    | Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.       |        |           |



QY 1410 caagtcacgttaactgctggtgatacttaagaaccccttctcaggtacacaggtacaccccc 1469  
 |||||||  
 Db 64 CAAATCCAGTAACCTGCTGATCTAAAGAACCCCTTCTTCAGGTACACAGGTACAAACCCC 123  
 QY 1470 atcaacccccacgtggtcacaactacagctctccctcgagaatgatgtggaacacaggaag 1529  
 |||||||  
 Db 124 ATCAACCCCTACTGCTGCTACACTACAGCTCTCCCTCGGAAATATGTGAAACACAGGAAG 183  
 QY 1530 caccataatctcagcagcagcgggtgtgctgtgctgaaatgctcactgctcactagacactgag 1589  
 |||||||  
 Db 184 CACCTATATATCTCAGACACGGGGGTGCTGTGGCTGGAATGCTACTGCTACTAGACTGAG 243  
 QY 1590 caggttatctgcggcggtgctcaggtgtgtgtgtacaaacactgattcccttagtaaacac 1649  
 |||||||  
 Db 244 CAGGTATATCTCCGCGGCTCCAGTGTGGTCAACAACCTGATTCCTTAGCTAACAC 303  
 QY 1650 agggattgtcaatcacacacacactcccgatggtgcataatagaacagcaggtattgtca 1709  
 |||||||  
 Db 304 AGGATTTGTCATACACACCCACTCCCGATGGCTCATATAGACACGCGCATTTGTCCA 363  
 QY 1710 aggtccctcaggtgcccagggagcggcgtgagtcagtcaggtgccactatgacagtaacaa 1769  
 |||||||  
 Db 364 AGGTCCTCAGGTGCCGAGGAGGGGGGTAGTCCAGTGGCCACTATGCACTACAA 423  
 QY 1770 ccaattcaactggtgtggtgctcactctatagtcctgcgcctgcccattgtccacgacaa 1829  
 |||||||  
 Db 424 CCAATTCACCAATGGGAGGCTGCTGCTATGAGCTTCGCCCATGTCATCCGACCAA 483  
 QY 1830 caccatgcacatgaggtgtggtgtgctcagcagcagcagcagcagcagcagcagcagcagc 1889  
 |||||||  
 Db 484 CACCATGACATATGGAGATTAGTGCTCTCCAGCCGACGACAGCTGCGACAGCAAGCAC 543  
 QY 1890 caggaaacccaatcaatgaatcagcagcagcagcagcagcagcagcagcagcagcagcagc 1949  
 |||||||  
 Db 544 CAGAAACCAATATCAAGTCCAGCGCCGACAGCAGTGGCTGCCCTTCTGTA 603  
 QY 1950 gaagtgttgaacacccggtacag-cctcctgtctatgtggaactgtgacaaatttctac 2008  
 |||||||  
 Db 604 GAAAT--GTGAACACCCGGTACAGCCCTCTTGTATGGGAACTTGGACATTTGTGAC 661  
 QY 2009 ac 2010  
 ||  
 Db 662 AC 663

RESULT 6  
 BF608742 688 bp mRNA EST 01-APR-2001  
 LOCUS M11\_001664 Mouse 9-day fetus cdna library ICRP522 Mus musculus  
 DEFINITION  
 CNA clone ICRP522P146 5', mRNA sequence.  
 ACCESSION BF608742  
 VERSION BF608742.1 GI:13505234  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Yabuyavi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B. G.,  
 Lehrich, H. and O'Brien, J.  
 Detection of a high number of novel genes in a 9-day mouse embryo  
 cDNA library normalised by oligonucleotide fingerprinting  
 Unpublished (2001)  
 CONTACT: Hennig S  
 Laboratory 123, dept. Lehnach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihnestr. 63-73, D-14195 Berlin, Germany  
 Tel: +49 30 8413 1612  
 Fax: +49 30 8413 1380  
 Email: hennig@molgen.mpg.de  
 EST's are made from clones being representatives of clone clusters.  
 Clone clusters were calculated from oligonucleotide fingerprints.  
 PCR Primers

FEATURES  
 source  
 1..688  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="ICRP522P1464"  
 /clone\_lib="Mouse 9-day fetus cdna library ICRP522"  
 /tissue\_type="whole embryo"  
 /dev\_stage="embryonic 9-day"  
 /lab\_host="E.coli, XL1 blue"  
 /note="Vector: pSVSPORT1; site.1: NotI; site.2: SalI;  
 library preparation by oligo dT priming of RNA. Clones can  
 be ordered from the Resource Center in Berlin,  
 http://www.rzpd.de."

BASE COUNT 150 a 240 c 147 g 142 t 9 others  
 ORIGIN

Query Match 19.5%; Score 610.6; DB 151; Length 688;  
 Best Local Similarity 95.7%; Pred. No. 7.2e-154;  
 Matches 644; Conservative 0; Mismatches 27; Indels 2; Gaps 2;

QY 2050 tggcccgagcgcgtgtgtgctgcatacttacaatagatcctgttgggagcagccctcat 2109  
 |||||||  
 Db 16 TTGGTCCGAGCGGGTGTGCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 75  
 QY 2110 cctgttctgtaactctcactctgactctgactctgtgacatctgtctggaagagcagttctc 2169  
 |||||||  
 Db 76 CCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 135  
 QY 2170 ccccaacccccacagctgtgacactgacagcagcagcagcagcagcagcagcagcagcagc 2229  
 |||||||  
 Db 136 CCCCAACCCCAACAGCTCCACCTGACCAAGCCGAGAGAGCCAGCACTTCCACACACAGA 195  
 QY 2230 cctggcagcaccacaccccaacacccgctcctgtgacactccctcactcagctgggtgtgacgac 2289  
 |||||||  
 Db 196 CCGGAGAGACCCACCCCAACCCGCTGCTGACCTCCCTCACTGAGGCTGGGCTGGACAC 255  
 QY 2290 agcagctgagcctctcctcactcaactacagcagcagcagcagcagcagcagcagcagcagc 2349  
 |||||||  
 Db 256 AGCCGCTGACCTCTCTCTCAACTACAGGCCACATGTCACACATGAGCGGCTGACATGCT 315  
 QY 2350 gcttttttaatttaatttttttttttttttttttttttttttttttttttttttttttt 2409  
 |||||||  
 Db 316 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 375  
 QY 2410 aaaccgcgtgctgcgc-gccaagcagcagcagcagcagcagcagcagcagcagcagcagc 2468  
 |||||||  
 Db 376 AAACCGGCTGGCGCTGGCCAGAGCCAGCATCCCTGTTCTCAGGCCAGAGGTTCTAGGTT 435  
 QY 2469 aggggtggtcctgtcaagcctcagagtggtgacagc-cctccaccaaagggttccac 2527  
 |||||||  
 Db 436 AGGGGTGGCCCTTCAAGCCTTCAAGCTGAGGTGGGACACACCCCTCCACCAAGGGTTCCAC 495  
 QY 2528 ttaacttgaatgtaacaaacacacacacagcgtgtcaaaagcgtatgctcctacttctgtac 2587  
 |||||||  
 Db 496 TCAAACTTGAATGTAACAAACCAACCCAGCTTTTCCAAAGGCTTAATCCCTTACTTCTCTAC 555  
 QY 2588 tctcctgtcctagccctaaagccccctccatcaaaagcgtgaaacagcagcagcagcag 2647  
 |||||||  
 Db 556 TGTCTGCTGCTGAGCCCTGAATGCCCCCTTCATCACTGAAGCTTGAACAGCAGCCACAGAG 615  
 QY 2648 tgttcaacctgtggtactgtggcagcagcagcagcagcagcagcagcagcagcagcagcagc 2707  
 |||||||  
 Db 616 TGTGTACACCTGGCTACTGCGGAGAGANAGANACTCAAGATGTGTACACACACACA 675  
 QY 2708 aggaagcgctcct 2720  
 |||||||  
 Db 676 AGGAAGGCGCTCT 688







RESULT 10  
 BG420290 919 bp mRNA EST 14-MAR-2001  
 LOCUS 602448309f1 NIH\_MGC\_14 Homo sapiens cDNA clone IMAGE:4586763 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG420290  
 VERSION BG420290.1 GI:13326796  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 919)  
 NIH-MGC <http://mhc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: DCTD/DRP  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM1316 row: a column: 04  
 High quality sequence stop: 661.  
 Location/Qualifiers  
 1. 919  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4586763"  
 /clone\_1lb="NIH\_MGC\_14"  
 /tissue="renal cell adenocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: kidney; Vector: pOTB7; Site.1: XhoI; Site.2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 176 a 281 c 269 g 193 t  
 ORIGIN

Query Match 17.8%; Score 555.2; DB 153; Length 919;  
 Best Local Similarity 85.5%; Pred. No. 7.5e-139;  
 Matches 677; Conservative 0; Mismatches 108; Indels 7; Gaps 5;

691 gaagaagctcgtgtaagaagtaacaatcgaacgacgacgtgctcaccctggaa 750  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 3 GCTGAGGCTTGGTGAAGATACCACTGACGACCGCATCTGGTCAATCCGGGCAAG 62  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 751 gaagaagagctcattgctgacgaagtgagacatatactcagaagccatgggtac 810  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 63 GGGAGAGAGGTGTCACCTCCCGAGAGGTGACATCATCTCGAGCCCATGGGCTAC 122  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 811 atgctcttaagaagaatgctcgagagctaccctcacaagcaaaagtacctgaagct 870  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 123 ATGCTCTTCAACGAGCGCATCTGGAGAGCTACTCCACCCCAAGAGTACTGAAACCC 182  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 871 agtggaaacaatctccaccattgtagatgacacgtcaccctcactgataagacag 930  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 183 AGGGAAACATGTTTCCATACATTGGTGAAGTCCACCTTCACCTTCACGAGTGAACAG 242  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 931 ctctaatagagagctgacccaagccaactccggtaccagccatcctcctcagagtg 990  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 243 CTCTAATAGAGAGTTCACCAAGGCCAACTTCTGTAACAGGCACTTTCACATGAGGTG 302

991 gacctgcggccctcagaaggtgcccgtgtgaaagtaagtaacttcggcaaccctgtgtgac 1050  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 303 GACCTGTGGCCCTCCAGAGGTGCCGCTGATAGTATTTCCGGACCTGTGTGGAC 362  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 1051 acatttacaatccggaatcctgatatggccaatctgtcaagtaacagaagtaactttagaa 1110  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 363 ACAATTACATCCGGATCCGATGCGGCAAGTCTCAAGTACAGGCTGAACCTTTAGAA 422  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 1111 gccaaagaagcgatttgcacagagatagaatcccatccaatccacaatctgattca 1170  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 423 GCCAAMAAGAGATTTGCCAAGATAGAAATCCCATTCAAATTCACATCTGCATTCA 482  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 1171 gggctagcatgattggtgaccttcgtgttgatgttgcatttcattggtcccaatgacc 1230  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 483 GGGCTGTGCCACGGC -TGGCTTTTGGTTGACCTTGTTCATTCGCTCCATATAGAC 541  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 1231 gtgtgataccacagcccacacagagccctgacccactgtatccaggttcggtgctc 1290  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 542 GTGTGGCTGTCCACAGGCCGACAGAGCCCTGACCCACTGTGACAGGTGC -GGTGTG 599  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 1291 ttccagtcaccgtgtgttgcacaagggccgggaacagctctcaaggaacatgtctctatt 1350  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 600 TTCCAGTCCACACCTGTTCGCC -AGGCAAGGGGACACGCTCTCAGGACATGCTTGCTTAT 658  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 1351 gccacaagaagacagagctatgacatgattgtgtgcacaggtggaccagacaggtcc 1410  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 659 TGCACAGAGACAGACAGACTTACGACTCCGTTTGTGC -CCGTTGGGGCGACCGCGCTCC 717  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 1411 aagtcagtaactgtctgataaagaacccctcttcagatgatacagaagtaacaccca 1470  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 718 AAGTCC -GTACCTCTGGATGTGAAGGCCCTTGTGATTACGGGCAAGGCTCACCCCG 775  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 1471 tcaccccaact 1482  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 776 CCGCTCCTACT 787  
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 11  
 BG063675/c 578 bp mRNA EST 26-JAN-2001  
 LOCUS H3010B08-3 NIA Mouse 15K cDNA clone set Mus musculus cDNA clone  
 DEFINITION H3010B08 3', mRNA sequence.  
 ACCESSION BG063675  
 VERSION BG063675.1 GI:12546238  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 578)  
 Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka  
 'T.S., Carter, M.G. and Ko, M.S.H.  
 TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set  
 JOURNAL Unpublished (2001)  
 COMMENT Other ESTs: H3010B08-5  
 Contact: George J. Kargul  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: [cdna@lsun.grc.nia.nih.gov](mailto:cdna@lsun.grc.nia.nih.gov)  
 This clone set has been freely distributed to the community. Please  
 visit <http://lsun.grc.nia.nih.gov/cDNA/15k.html> for details.  
 Seq primer: -21M13 Forward  
 Plate: H3010 row: B column: 08  
 High quality sequence stop: 578  
 POLYA=yes.  
 Location/Qualifiers  
 1. 578  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="H3010B08"  
 /clone\_1lb="NIA Mouse 15K cDNA clone set"





```

1. .584
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U1-M-A10-aax-f-06-0-01"
/clone_lib="NIH BMAP MOB"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: p773d-pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH BMAP MOB library is a non-normalized library
constructed from mouse olfactory bulbs. The tag is a
string of 5 nucleotides present between the Not I site

```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Mus  
1 (bases 1 to 713)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/cgap>.  
National Cancer Institute, Cancer genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2001, 18:45:25 ; Search time 89.36 Seconds  
(without alignments)  
6476.648 Million cell updates/sec

Title: US-09-464-377-1

Perfect score: 3124  
Sequence: 1 agggggcctgagcgagcggacc.....taaaagtgttcttgytaa 3124

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 317530 seqs, 92630169 residues

Total number of hits satisfying chosen parameters: 635060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCITUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Query Length | DB ID | Description                             |
|------------|--------|-------------|--------------|-------|-----------------------------------------|
| 1          | 1494.2 | 47.8        | 2224         | 3     | US-09-109-204-6<br>Sequence 6, Appl     |
| 2          | 493    | 15.8        | 621          | 3     | US-09-109-204-24<br>Sequence 24, Appl   |
| 3          | 443    | 14.2        | 549          | 3     | US-09-109-204-25<br>Sequence 25, Appl   |
| 4          | 420    | 13.4        | 647          | 3     | US-09-109-204-26<br>Sequence 26, Appl   |
| 5          | 284.2  | 9.1         | 631          | 3     | US-09-109-204-23<br>Sequence 23, Appl   |
| 6          | 222.2  | 7.1         | 258          | 3     | US-09-109-204-22<br>Sequence 22, Appl   |
| 7          | 221.4  | 7.1         | 256          | 3     | US-09-109-204-21<br>Sequence 21, Appl   |
| 8          | 139.8  | 4.5         | 529          | 3     | US-09-109-204-28<br>Sequence 28, Appl   |
| 9          | 107.8  | 3.5         | 220          | 3     | US-09-109-204-29<br>Sequence 29, Appl   |
| 10         | 106.6  | 3.4         | 655          | 3     | US-09-109-204-27<br>Sequence 27, Appl   |
| 11         | 70.2   | 2.2         | 579          | 4     | US-08-991-789A-147<br>Sequence 147, App |
| 12         | 48.6   | 1.6         | 1028         | 4     | US-08-118-200-1<br>Sequence 1, Appl     |
| 13         | 48.6   | 1.6         | 1028         | 4     | US-08-458-745-1<br>Sequence 1, Appl     |
| 14         | 48.4   | 1.5         | 4524         | 2     | US-08-845-998-7<br>Sequence 7, Appl     |
| 15         | 48.4   | 1.5         | 4524         | 2     | US-09-206-537-7<br>Sequence 7, Appl     |
| 16         | 46.6   | 1.5         | 7218         | 1     | US-08-232-463-14<br>Sequence 14, Appl   |
| 17         | 45.4   | 1.5         | 3765         | 4     | US-07-705-490-1<br>Sequence 1, Appl     |
| 18         | 45.4   | 1.5         | 3765         | 4     | US-07-751-891B-1<br>Sequence 1, Appl    |
| 19         | 45.4   | 1.5         | 4362         | 2     | US-08-455-073A-1<br>Sequence 1, Appl    |
| 20         | 43.4   | 1.4         | 530          | 4     | US-08-758-662-4<br>Sequence 4, Appl     |
| 21         | 43.4   | 1.4         | 4480         | 4     | US-09-191-171-7<br>Sequence 7, Appl     |
| 22         | 43.4   | 1.4         | 4480         | 4     | US-09-385-707-7<br>Sequence 7, Appl     |
| 23         | 43     | 1.4         | 8438         | 1     | US-07-945-283-1<br>Sequence 1, Appl     |
| 24         | 42.4   | 1.4         | 2150         | 2     | US-08-318-837-1<br>Sequence 1, Appl     |
| 25         | 42.2   | 1.4         | 503          | 4     | US-08-483-533-7<br>Sequence 7, Appl     |
| 26         | 42.2   | 1.4         | 1292         | 4     | US-08-483-533-37<br>Sequence 37, Appl   |
| 27         | 42.2   | 1.4         | 152331       | 4     | US-09-128-155-16<br>Sequence 16, Appl   |

|    |      |     |       |   |                                       |
|----|------|-----|-------|---|---------------------------------------|
| 28 | 42   | 1.3 | 2991  | 4 | US-08-795-430-48<br>Sequence 48, Appl |
| 29 | 41.8 | 1.3 | 14060 | 3 | US-08-658-136-4<br>Sequence 4, Appl   |
| 30 | 41.6 | 1.3 | 2467  | 3 | US-08-753-007A-1<br>Sequence 1, Appl  |
| 31 | 41.6 | 1.3 | 2467  | 4 | US-09-398-496-1<br>Sequence 1, Appl   |
| 32 | 41.4 | 1.3 | 3472  | 6 | 5244792-1<br>Patent No. 5244792       |
| 33 | 41.1 | 1.3 | 936   | 1 | US-08-018-977C-4<br>Sequence 4, Appl  |
| 34 | 40.8 | 1.3 | 1650  | 2 | US-08-776-210-6<br>Sequence 6, Appl   |
| 35 | 40.6 | 1.3 | 44377 | 2 | US-08-804-227C-7<br>Sequence 7, Appl  |
| 36 | 40.6 | 1.3 | 44377 | 2 | US-08-804-198-1<br>Sequence 1, Appl   |
| 37 | 40.2 | 1.3 | 6085  | 4 | US-09-029-603-4<br>Sequence 4, Appl   |
| 38 | 40   | 1.3 | 18994 | 1 | US-08-459-586-4<br>Sequence 4, Appl   |
| 39 | 40   | 1.3 | 18994 | 2 | US-08-282-696-4<br>Sequence 4, Appl   |
| 40 | 39.8 | 1.3 | 1776  | 1 | US-08-722-001-29<br>Sequence 29, Appl |
| 41 | 39.8 | 1.3 | 2140  | 1 | US-08-334-698-1<br>Sequence 1, Appl   |
| 42 | 39.8 | 1.3 | 2140  | 1 | US-08-228-932-1<br>Sequence 1, Appl   |
| 43 | 39.8 | 1.3 | 2140  | 1 | US-08-468-939-1<br>Sequence 1, Appl   |
| 44 | 39.8 | 1.3 | 2140  | 2 | US-08-406-855A-1<br>Sequence 1, Appl  |
| 45 | 39.8 | 1.3 | 2140  | 3 | PCT-US95-04203-1<br>Sequence 1, Appl  |

#### ALIGNMENTS

RESULT 1  
US-09-109-204-6  
Sequence 6, Application US/09109204  
Patent No. 6060250  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Gueglere, Karl J.  
APPLICANT: Gorgone, Gina A.  
APPLICANT: Corley, Neil C.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: HUMAN TRANSFERASES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/109,204  
FILING DATE: HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
FILING DATE:  
NAME: Cerrone, Michael C  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0546 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-855-0572  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2224 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRAITUT21  
CLONE: 2525071

US-09-109-204-6

|                       |              |               |                 |              |
|-----------------------|--------------|---------------|-----------------|--------------|
| Query Match           | 47.8%;       | Score 1494.2; | DB 3;           | Length 2224; |
| Best Local Similarity | 82.8%;       | Pred. No. 0;  |                 |              |
| Matches 1896;         | Conservative | 0;            | Mismatches 308; | Indels 87;   |
|                       |              |               |                 | Gaps 13;     |

|    |      |                                                                    |      |
|----|------|--------------------------------------------------------------------|------|
| OY | 290  | cagggtcagtcgtgtgtggagagcaatctctcaatacaaccttgggtctgcaacagcgcc       | 34   |
| Db | 1    | CAGAGTCGACCGCTGTGTGGGACGACAGTCTTCAATCAATCACTGGGCTGCAACACGCTCC      | 60   |
| OY | 350  | tcatcagatttgcacacacccacagatt--tcgtctcttctcaacaatccttgaacctg        | 407  |
| Db | 61   | TCATCCAGTTTGGCACACACCAAGATTTCGGCTGAGCTTCTFACAAATCTGAACACCTG        | 120  |
| OY | 408  | tcgggtgccacaacttggtagcgctctctgttctcagttagcggtgacagaggaatcctcagctgt | 467  |
| Db | 121  | CCGGGGGCCACACCTTGGAGGGTGTGTGTTCAACGAGCGAGGGAGGAGTCTTTCGCCGT        | 180  |
| OY | 468  | gcagattctcaagttctctatgtctacctaaccacagagcaacatgatacagactatgt        | 527  |
| Db | 181  | GCACTACTTCCAGTTTATATGCTACCTGTCCAGCAGCAAGAACATGATACAGACTACGT        | 240  |
| OY | 528  | gcggacagagcaactcaacagcgltgcgtcctgcgaaacacacagyaacttaagyaacagt      | 587  |
| Db | 241  | GGGAGCAGGACACTACCAAGCGGCCCATCTCGCAAAACACACCGACTTCAAGACAAAT         | 300  |
| OY | 588  | gattctagaatgtggcgtgtggctctggatccctgtacatttttggcttcaagcaagsgc       | 647  |
| Db | 301  | CGTCTGTGATGTGTGGCTGTGTGGCTCTGGGATCCGTGTGTTTTGGCGCCCAAGCTGGAGC      | 360  |
| OY | 648  | caggaaattatltcagtgtggaagccagacacatgtgctcagcatgcaagagttcgtgtaa      | 707  |
| Db | 361  | ACGGAAATCTACGCGGTGGAGGCGACACACATAGGCCACGACGCTGAGGTGTGTGA           | 420  |
| OY | 708  | gaatfaacaatctgcacagacgcgatcgtgtgtcatccctgtgcagaagttagaggttcat      | 767  |
| Db | 421  | GAGTAAACAACCTGACGGACCGCATCTGTGTGATCCCGGGCAAGGAGGAGAGTGTACT         | 480  |
| OY | 768  | gcctgagcaagtltgacaattatcatctcaagggccatggtgtacatgcttctcaagyaag      | 827  |
| Db | 481  | CCCCGACGAGTGGACATCTCATCTTCGAGGCCATGGGCTACATGCTTTCAACGACG           | 540  |
| OY | 828  | aatgctcgagagctacctctcatgccaanaatgtaacttgaagcctagtgtgaacaatgtccc    | 887  |
| Db | 541  | CATGCTGGAGACTACTCCAGCCCAAGAAAGTACGTGAAGCCAGGCAAGCAATGTTCC          | 600  |
| OY | 888  | caccattgtgtatgttcaactctgcaccccttcaactatgttgaagcgtctcatgtgaacatt    | 947  |
| Db | 601  | TACATTGTGTGACCTCCACCTTCACCTTCACGGATGAACAGCTCTATCATGGAACAATT        | 660  |
| OY | 948  | caccaaaagccaacttccggtgacagacatcctctcaatgttagttagtgcactgtgcgacctag  | 1007 |
| Db | 661  | CACCAAGGCCAACTTTCGTAACGACGACATTTCCATGAGATGTGACGTGCGCCCTCG          | 720  |
| OY | 1008 | aggtgtccgctgtgtgtagtgaacttcggcaactgtgtgtgtgagacatttgaatccgagt      | 1067 |
| Db | 721  | AGGTGCGCGGTGATAGTAATTTCCGGCACGCTGTGTGTGACACATTTGACATCCGAT          | 780  |
| OY | 1068 | ccttgaatgcacaaactctgcaagtaacagatgaactccttagaagccaagaagcgcaatt      | 1127 |
| Db | 781  | CTGTATGGCCAAAGTCTGTCAAGTACAGGGGAACCTTTAGAAAGCAAGAGGAGATT           | 840  |
| OY | 1128 | gcacaagatagaaatcccatccaatccaatctgcatagtctgaatcaaggtctagtcatgtctt   | 1187 |
| Db | 841  | GCAAGAGTAGAAATCCCATTTCAAAATTCACATGCGATTCAGAGGCTGTGTCCAGGGCT        | 900  |
| OY | 1188 | gaccttcaggttgaatgttgccttcaatgtgtcacaatgagacgtgtgtgtctaccagc        | 1247 |
| Db | 901  | GGCTTTCTGGTTTACGTGTGCTTTCAATCGGCTTCATATATGACCGTGTGCTTCCACAGC       | 960  |
| OY | 1248 | cccaacagagcccttgcacccactgttacaagatgcggtgtcctcttcacgtacagtglt       | 1307 |

|    |      |                                                                   |      |
|----|------|-------------------------------------------------------------------|------|
| Db | 961  | CCGACAGAGCCCTCAGACCCACTGGATACAGAGTGGCGTGGCTGTTCACAGTACCACTGTT     | 1020 |
| Qy | 1308 | tgccaagcgccggggagacaagctctcaaggacatgctctgtattatgcaacaagaagacag    | 1367 |
| Db | 1021 | CGCAAGGACAGGGGACACGCTCTCAGGGACATGTGCTATTGSCAAACAAAGACAGAG         | 1080 |
| Qy | 1368 | ctatgaatcaatgatttgygacaggttggacaaagacagctccaagttccaagtccaagtccgtc | 1427 |
| Db | 1081 | CTACGCAATCAGATATGTGGCCAGGTGGACCCAGACCGGTCTCCAGTCCAGTAACTCCT       | 1140 |
| Qy | 1428 | ggaatcaagaacccctttcttaagfaaacaggttaaaccccaatacccccaactgctgc       | 1487 |
| Db | 1141 | GGATCTGAAAAACCCCTTCTTTAGATACAGGGCACACCCCTCACCCACCCCGGCTC          | 1200 |
| Qy | 1488 | acaactacaagcttcctctcgagagatatgtgnaacacagaagaacactataatctcaag      | 1547 |
| Db | 1201 | CCACTACATCTCCTCCTCGGAAACATGTGGAACACGGGAGAGCACTACAACTCAGAG         | 1260 |
| Qy | 1548 | cggggtgagctgtggtcgtgaattgctaactgtctaagacactgaagcagtgttatgtccggcg  | 1607 |
| Db | 1261 | CGGATGCGCGTGGCGAGGATGCCGACCCGCTATGACTTGACAGAGTGTATTGCGCAATGG      | 1320 |
| Qy | 1608 | ctccaagtgtggtgaacaaacaaactgattcccttagtataacagggatgttcaatcac       | 1667 |
| Db | 1321 | CTTCACAGGTGGGCGACAAACACTGTATCCTTTAGCCAAACAGGGGATTTCTCAATCAC       | 1380 |
| Qy | 1668 | ccaactccggaatggagctcacaataatgacacagggacatgtccaaagcctccagtgcca     | 1727 |
| Db | 1381 | CCACTGCCGATGGGCTCCATATATGACAGGGGATTTGTCCAAAGGTCCTCCGGCGCCA        | 1440 |
| Qy | 1728 | gg--gaggcgcggtagctccaagtcagcaatgccaatgcaagtaacaacaaagttcaacatg    | 1784 |
| Db | 1441 | GGGCAATGTGTGGGACAGACAGAGTCCACATATGACATCAACAGCGCATGTTCACATAGG      | 1500 |
| Qy | 1785 | tggcctctgcaatctctatgagctcgcccaatgttcaatccgaaccaacaacatgtcaatg     | 1844 |
| Db | 1501 | CGGCGCGGCATCTCCATGAGCGTGGCCCATGTCCATCCACCAACCACTATGACATAGG        | 1560 |
| Qy | 1845 | gaatgaagtgtgcctccagcgcgacagcaactgtgcgactgaacagcaccaagaaaccaatca   | 1904 |
| Db | 1561 | GAGCTAAGGGGC-----CGCGCCCGCGGACTGACAGCACAGAAACCAATGA               | 1608 |
| Qy | 1905 | agtcacagc---ccggacaaagccagtgagctgtttcccccctgtctctggaagaattgtgtaa  | 1961 |
| Db | 1609 | TGTGCTGCGCGCGCGCCCGCGCGGGGGGTTCGCCCTGTGTAAGGAGAAG--CTCGAA         | 1666 |
| Qy | 1962 | caaccggtcaacagc-ctccttgctatgggaacttggacaatttgttcaacagatgtgcgcg    | 2020 |
| Db | 1667 | CACCCCGTACAGCTCTCTTGTGTAGGGAACTGGGACACTTTTATTAACAGATGTTCCG        | 1726 |
| Qy | 2021 | ctgcacctaaagtaaccccccaaacctttgtgtgcgagcggtgtgtgtccatactt          | 2080 |
| Db | 1727 | CCGTCCCCA---CCCTAACCCCACTCCCGGCGCTGAGCGTGTGTGCGCATATT             | 1783 |
| Qy | 2081 | acaatgaatcctgttggggcagcctcat--cctgtctgttactctccactctgaacctg       | 2138 |
| Db | 1784 | ACACAAATATCATGTGTGGAGCGCTCTGTCGCCCTCCCGCTACCTGACCTGACCTGG         | 1843 |
| Qy | 2139 | ctttgacatctgtctgaagaaggaagtctcctcccccaaccccccaagttgaacctgcacag    | 2198 |
| Db | 1844 | GCTGTGATCTGTGTGAACAGGGGCGCATGAGGGGCTCGACACCTGCTGTC-----           | 1893 |
| Qy | 2199 | gcaagagggagccagagctgcacacacagaactgcagcagaccacaccccaacccgtcc       | 2258 |
| Db | 1894 | -----CAGGTCCCTTATGACACTGTCC                                       | 1914 |
| Qy | 2259 | tgtcacctccctcacacttggtgtggaagaacagccagctgagacctctccttcaactaca     | 2318 |
| Db | 1915 | CCCTGCTGTCTCCATGTGGAAAGTGTAGCTGCGCAGGCGGGCGCTCCCTTCGACAGCA        | 1974 |
| Qy | 2319 | gagcaaatgtgtcaacatgggcgtgaaatgctgttttttaattattatttttttaagaaa      | 2378 |
| Db | 1975 | GGCTCT--GGTACAAACGGAACGTGATGCTGTTTTTTTAAATTTATTTTATGAAA           | 2032 |









FILED DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cerrone, Michael C  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0546 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-855-0572  
TELEX:  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 258 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BLADTUT07  
CLONE: 1889292H1  
US-09-109-204-22

Query Match  
Best Local Similarity 7.1%; Score 222.2; DB 3; Length 258;  
Matches 233; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 290 cagagtgacgtgctgtgagcagacagtccttcacacccctggcgtgacagcgtcc 349  
DB 1 CAGAGTGCAGCCCTGGGCGAGCAGTTCCTTCATCACCCCTGGCGTCCAGCGCTCC 60  
QY 350 tcatcagtttgcacacccacagatttctgtctttctaacacaccttaaacctgtc 409  
DB 61 TCATCAGTTTGCACACCCAGCAGTTCCTTCATCACCCCTGGCGTCCAGCGCTCC 120  
QY 410 gggagcaccactgagcgcctgtgtttcagtgagcagaggaatccctcagctgtgc 469  
DB 121 GGGGCGCACACCCCTGGGCGGTGTGTTCAGCAGCGGAGAGTCTTTCGCGCTGC 180  
QY 470 agtacttccagttctgtgtctactctaccagacagagaacatgagcagactatgtgc 529  
DB 181 AGTACTTCCAGTTTATGGCTACTCTGCCAGCAGAGAACATGATCAGAGACTACGTGC 240  
QY 530 ggaacagcact 541  
DB 241 GGACAGCACCT 252

RESULT 7  
US-09-109-204-21  
Sequence 21, Application US/09109204  
Patent No. 6060250  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
APPLICANT: Gorgone, Gina A.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: HUMAN TRANSFERASES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/109,204  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cerrone, Michael C  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0546 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-855-0572  
TELEX:  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 256 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRAITUT21  
CLONE: 2525071H1  
US-09-109-204-21

Query Match  
Best Local Similarity 7.1%; Score 221.4; DB 3; Length 256;  
Matches 234; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 385 ttctacaacatccctgaaacactgtcgagggccacacacacgtggagcgtctgtgttcagtgag 444  
DB 2 TTCTACAACATCCGAAACCTGCGGGGCCACACCCCTGGAGCGGTGTGTTCAGCGAG 61  
QY 445 cggacagaggaatccctcagctgtgtcagtaacttccagttctatgtgtactatccacag 504  
DB 62 CGGACGAGAGAGTCTTCTGCCGTGCAGTACTTCCAGTTTATGGCTACTCTCCAGCAG 121  
QY 505 cagaacatgtatgcagacatgtgtgcagacagacacacacacagcgtgcagatccctcagag 564  
DB 122 CAGAACATGATGACGAGTACTGCGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 181  
QY 565 cacaagacttcaagagacagatcgtctagatgttgagcgtctggagacactgtgc 624  
DB 182 CACACCGACTTCAAGAGACATGTTCTTGATGTGGCTGTGGCTGTGGATCTGTGC 241  
QY 625 ttlttgctgtcaca 639  
DB 242 TTTTTCGCCGCCAA 256

RESULT 8  
US-09-109-204-28  
Sequence 28, Application US/09109204  
Patent No. 6060250  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
APPLICANT: Gorgone, Gina A.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: HUMAN TRANSFERASES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:

MEDIM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: Windows  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/109,204  
 FILING DATE: HEREWITH  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ceirone, Michael C  
 REGISTRATION NUMBER: 39,132  
 REFERENCE/DOCKET NUMBER: PF-0546 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-855-0555  
 TELEFAX: 650-855-0572  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 28:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 529 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: BRSTN09  
 CLONE: 1664573F6  
 US-09-109-28

|                       |              |                   |               |            |
|-----------------------|--------------|-------------------|---------------|------------|
| Query Match           | 4.5%         | Score 139.8       | DB 3          | Length 529 |
| Best Local Similarity | 77.0%        | Pred. No. 1.9e-25 |               |            |
| Matches 218           | Conservative | 0                 | Mismatches 35 | Indels 10  |
|                       |              |                   |               | Gaps 4     |

|    |      |                                                              |      |
|----|------|--------------------------------------------------------------|------|
| OY | 2295 | gcctgacccctccctcaacacacagagccacatggtacacatggcggtgacatcgtctt  | 2355 |
| Db | 34   | GCGGGGCCCTCCCTTCACACACAGGCTTC--GGTACACAGGACCTGCATGCTGCTT     | 91   |
| OY | 2355 | ttttaattatttttttttcgaaagaacacagtgtaaccacacagacccctcgagaac    | 2414 |
| Db | 92   | TTTTAAATTTATTTTTATGAAAGAAGACAGTGTCATGCCGAGACCTCTGTGAAGC      | 151  |
| OY | 2415 | cggctgg-cgcgcgaagccagcccccgttctctatggcccagaag-ttctagtgaagg   | 2474 |
| Db | 152  | AGGCCGGCCGGGCGGACCCAGNCGCCCTCTCCCTGACTCAGAGGCGCGGGGAGGG      | 211  |
| OY | 2473 | gtggccctgtcaagcctctcagagtgycacagcccccctccacacaaagggttaactcaa | 2533 |
| Db | 212  | GTNNCCCGCGGAGGCTTCAGGG-----NNNNCCCTCCCAACCAAGGGTTCACTCA      | 265  |
| OY | 2533 | cttgaatgtataccaaccaacagctgtctcaaaaggcctagtccct               | 2575 |
| Db | 266  | CTTGAATGTACANCCANCCACTGTTCGGGAAGCTCCCTGCTCC                  | 308  |

RESULT 9  
 US-09-109-204-29  
 : Sequence 29, Application US/09109204  
 : Patent No. 6060250  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Lal, Preeti  
 : APPLICANT: Bandman, Olga  
 : APPLICANT: Hillman, Jennifer L.  
 : APPLICANT: Guegler, Karl J.  
 : APPLICANT: Gorgone, Gina A.  
 : APPLICANT: Corley, Neil C.  
 : APPLICANT: Patterson, Chandra  
 : TITLE OF INVENTION: HUMAN TRANSFERASES  
 :  
 : NUMBER OF SEQUENCES: 32  
 :  
 : CORRESPONDENCE ADDRESS:  
 :  
 : ADDRESSEE: Incyte Pharmaceuticals, Inc  
 : STREET: 3174 Porter Drive

CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: Windows  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/109,204  
 FILING DATE: HERewith  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ceirone, Michael C  
 REGISTRATION NUMBER: 39,132  
 REFERENCE/DOCKET NUMBER: PF-0546 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-855-0555  
 TELEFAX: 650-855-0572  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 29:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 220 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: LUNGTUT03  
 CLONE: 1474156T1  
 US-09-109-204-29

| Query Match           | 3.5%                                                            | Score 107.8     | DB 3     | Length 220 |
|-----------------------|-----------------------------------------------------------------|-----------------|----------|------------|
| Best Local Similarity | 78.9%                                                           | Pred. No. 1e-17 |          |            |
| Matches 168           | Conservative 0                                                  | Mismatches 37   | Indels 8 | Gaps 3     |
| QY 2357               | ttaatttattttttttaagaagaagaccagtgtaaccacagaccctctgagaaccg        |                 |          | 2416       |
| DB 6                  | TTAATTTATTTATTTTATGAAAAGAACCATGTCATTCGGAGACCTCTGTGAAGCCAG       |                 |          | 65         |
| QY 2417               | gctgagcgcgccaagccaagcagccctcttctctagggcccaagggttctaggtgaggggt   |                 |          | 2474       |
| DB 66                 | GCGCGCGCGGCGCGAGCCGAGACGCCCTCTCCCTTAACATCAGAGGGCGCGGGGAGGGGT    |                 |          | 125        |
| QY 2475               | ggccctgtccaagcctctcaagatggtgacacagccctccccaacaaaggtctacactcaaat |                 |          | 2534       |
| DB 126                | GGCCCCCGCGAGGGCTTCAGGG-----GCCCTCTCCCAACCAAGGGTTCACCTACACT      |                 |          | 179        |
| QY 2535               | tgaatgtacaaacccaagcagctgttccaagacc                              |                 |          | 2567       |
| DB 180                | TGAATGTACAAACCCACCCACTGTGGGGAGAGCC                              |                 |          | 212        |

RESULT 10  
US-09-109-204-27  
Sequence 27, Application US/09109204  
Patent No. 6060250  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Gorgone, Gina A.  
APPLICANT: Corley, Neil C.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: HUMAN TRANSFRASES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.





|                           |        |                  |           |              |
|---------------------------|--------|------------------|-----------|--------------|
| Query Match               | 1.58;  | Score 48.4;      | DB 4;     | Length 4524; |
| Best Local Similarity     | 53.58; | Pred. No. 0.013; |           |              |
| Matches 123; Conservative | 0;     | Mismatches 106;  | Indels 1; | Gaps 1;      |



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2001, 16:51:39 ; Search time 4073.48 Seconds

(without alignments)  
11862.400 Million cell updates/sec

Title: US-09-464-377-1

Sequence: 3124  
1 agggggcctgagcgacc.....taaaagtgttcttctgtaa 3124

Scoring table: IDENTITY\_NOC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_da1:\*  
2: gb\_da2:\*  
3: gb\_da3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_rod:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_vl:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_vl1:\*  
59: gb\_vl2:\*  
60: gb\_hrg1:\*  
61: gb\_hrg2:\*  
62: gb\_hrg3:\*  
63: gb\_hrg4:\*  
64: gb\_hrg5:\*  
65: gb\_hrg6:\*  
66: gb\_hrg7:\*  
67: gb\_hrg8:\*  
68: gb\_hrg9:\*  
69: gb\_hrg10:\*  
70: gb\_hrg11:\*  
71: gb\_hrg12:\*  
72: gb\_hrg13:\*  
73: gb\_hrg14:\*  
74: gb\_hrg15:\*  
75: gb\_hrg16:\*  
76: gb\_hrg17:\*  
77: gb\_hrg18:\*  
78: gb\_hrg19:\*  
79: gb\_hrg20:\*  
80: gb\_hrg21:\*  
81: gb\_hrg22:\*  
82: gb\_hrg23:\*  
83: gb\_hrg24:\*  
84: gb\_hrg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_r01:\*  
95: gb\_r02:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_da3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description         |
|------------|--------|-------------|--------|-------|---------------------|
| 1          | 3124   | 100.0       | 3124   | 94    | AF117887 Mus muscu  |
| 2          | 2256.4 | 72.2        | 2336   | 94    | BC003964 Mus muscu  |
| 3          | 1355.4 | 43.4        | 233830 | 75    | AC073731 Mus muscu  |
| 4          | 1141.2 | 36.5        | 2160   | 88    | AF055027 Homo sapi  |
| 5          | 745.8  | 23.9        | 1917   | 94    | BC003289 Mus muscu  |
| 6          | 745.8  | 23.9        | 2107   | 94    | BC002282 Mus muscu  |
| 7          | 409.2  | 13.1        | 500    | 10    | AX079470 Sequence   |
| 8          | 408    | 13.1        | 11218  | 63    | AC013033 Drosophila |



| RESULT                                                                                 | 1                                                                                                                                                       | ALIGNMENTS |
|----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| LOCUS AF117887                                                                         |                                                                                                                                                         |            |
| DEFINITION Mus musculus protein arginine methyltransferase (Carm1) mRNA, complete cds. |                                                                                                                                                         |            |
| ACCESSION AF117887                                                                     |                                                                                                                                                         |            |
| VERSION AF117887.1                                                                     | GI:5257220                                                                                                                                              |            |
| KEYWORDS                                                                               |                                                                                                                                                         |            |
| SOURCE                                                                                 | house mouse.                                                                                                                                            |            |
| ORGANISM                                                                               | Mus musculus                                                                                                                                            |            |
| REFERENCE                                                                              | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 3124) |            |
| AUTHORS                                                                                | Chen D., Ma H., Hong H., Koh S.-S., Huang S.M., Schurter B.T., Aswad D.W. and Stallcup M.R.                                                             |            |
| TITLE                                                                                  | Regulation of transcription by a protein methyltransferase                                                                                              |            |
| JOURNAL                                                                                | Science 284 (5423), 2174-2177 (1999)                                                                                                                    |            |
| MEDLINE                                                                                | 99316081                                                                                                                                                |            |
| REFERENCE                                                                              | 2. (bases 1 to 3124)                                                                                                                                    |            |
| AUTHORS                                                                                | Chen D., Ma H., Hong H., Koh S.-S., Huang S., Schurter B.T., Aswad D.W. and Stallcup M.R.                                                               |            |
| TITLE                                                                                  | Direct Submission                                                                                                                                       |            |
| JOURNAL                                                                                | Submitted (05-JAN-1999) Pathology, University of Southern California, School of Medicine, 2011 Zonal Ave. HMR 301, Alhambra CA 91801, USA               |            |
| FEATURES                                                                               | location/Qualifiers                                                                                                                                     |            |
| SOURCE                                                                                 | 1..3124                                                                                                                                                 |            |
| gene                                                                                   | /organism="Mus musculus"                                                                                                                                |            |
|                                                                                        | /db_xref="taxon:10090"                                                                                                                                  |            |
|                                                                                        | 1..3124                                                                                                                                                 |            |
|                                                                                        | /gene="Carm1"                                                                                                                                           |            |

[illegible]

OY 721 acagaccgcatcgtgtgcatccctggcacaagtagagagctcattgctcgtgagcaatg 780  
DB 721 ACAGACCGCATCGTGTGCATCCCTGGCAAGTAGAGAGGCTCATTCCTGAGCAAGG 780  
OY 781 gacattcatctagagcccatggctacatgctcttcaatgaagaatgctgagagc 840  
DB 781 GACATTTCATCTCAGAGCCCATGGCTACATGCTTCATTAAGAAAGATGCTCAGAGC 840  
OY 841 tacctcatgcaaaaagattacacctgaagcttagtgaacatgctcccaacattgatat 900  
DB 841 TACCTCATGCAAAAAGATTACACCTTAGTGGAACATGTTCCCAACCATTTGGTAT 900  
OY 901 gtccacatcgacccctcactactgataagcagcttactatgagcagttcaaccaagccac 960  
DB 901 GTCCACCTCGACCCCTTCACTGATGAAACAGCTTACATGAGCAAGTTCCACCAAGCCAAC 960  
OY 961 ttcccgtaacagcaatcccttccatgtagtggactgtgcgcctcagaagtgccgctg 1020  
DB 961 TTCGGTACAGCCATCCTTCCATGAGTGGAGCTGTGCGCCCTCAGAGTGGCCTGTG 1020  
OY 1021 gatgagactctccgagccactgtgtggaacacattgacatccgagatctcattgagcaaa 1080  
DB 1021 GATGAGACTCTCCGAGCAACCTGTGTGGACACATTTGACATCCGATCTCATGGCCAAA 1080  
OY 1081 tctgtcaagtaacagtgaaattctttagaagccaagaagcagatttgacagagatagaa 1140  
DB 1081 TCTGTCAAGTAAACAGTGAATCTTGTAGAACCAAGAGCGATTGTCACAGGATAGAA 1140  
OY 1141 atcccatcaaatctccacatgctgcatcattcaaggtcattgcatgcttgcttctgttc 1200  
DB 1141 ATCCCATCAAAATTCACATGCTGCAATTCAGGGCTAGTCCATGGCTTGTGGTTC 1200  
OY 1201 gatgtgcttcatattgctccataatgacgtgtgtgtatccagagcccaagagagcc 1260  
DB 1201 GATGTGCTTTCANTGGCTTCATTAATGACCGTGTGGCTATCACAAGCCCAAGAGCCC 1260  
OY 1261 ctgaaccaactgtgtaacaggtcgcgtgctcttccagtaaccgtgtgttgcgaagcgcg 1320  
DB 1261 CTGACCCACTGTGTAACAGGTGCGGCTCTTCCAGTACACGTTGTTGCCAAGCGCGG 1320  
OY 1321 gaaacgtctccagaggaatgtctgtctaatgtgcaaaaagaagaagatgacatcagt 1380  
DB 1321 GAAACGCTCTCAGAGGAATGTGCTTAATGCAAAACAAACAGAGATGAGATCAGT 1380  
OY 1381 attgtgcaacaggtggaacagcagagctccaagtcagtaaccgtcgtgataaagaac 1440  
DB 1381 ATTGTGCAACAGGTGGAACAGAGCTCCAAAGTCCAGTAACCTGCTGATCTAAAGAAC 1440  
OY 1441 cccttcttcaagtaacaggttacaaccccaaccacccactggtctacaactacgtct 1500  
DB 1441 CCCTTCTTCAAGTAAACAGGTACAAACCCCAACCCCACTGGCTCACAATACAGCTCT 1500  
OY 1501 cccctggagagaaatgtggaacacaggaagaactataatctcagcagcgagggtgctgtg 1560  
DB 1501 CCTTCGAGAAATGTGGAACACAGGAAGCAACCTAATCTCAGACAGGGGTGGCTGTG 1560  
OY 1561 gctggaatgctactgctcctcagcagccctgagcagtgatttgcggcgcgctcagatgtgt 1620  
DB 1561 GCTGGAATGCTACTGCTCTCAGACTAGCAAGTATTATTGCGCGCGCTCCAGTGTGGGT 1620  
OY 1621 caacaacaactgattcccttagtataacagaggtatgtlcaatacaacacactcccgatg 1680  
DB 1621 CACAACAACCTGATTCCTTAGTAAACACAGGAGTGTCAATACACCACTCCCGGATG 1680  
OY 1681 ggtcacaataatgagcagggcattgtlccaaagctcctcaggtgcagggagagcgaggt 1740  
DB 1681 GGTCAATATATGAGCAGGGCATTGTCCAAGGCTCCTCAGGTGCCCGAGGAGCGGGGT 1740  
OY 1741 agctcagtgccactatgagtaacaaacagtaacacagttgagtggtcctgcatcttc 1800  
DB 1741 AGCTCAAGTCCCACTATGAGTCAACCAACAGTTCAACCAATGGGTGGCCCTGCATCTCT 1800

OY 1801 atggcctcgcccatgtlccatcccgacccaacacatgcaactatgtaggttagtgcctca 1860  
DB 1801 ATGGCCTCGCCCATGTlccatcccgacccaacacatgcaactatgtaggttagtgcctca 1860  
OY 1861 gccgcgaacagctgcgaactgacagcagcagaagaaacaaatcaagtccaagccggcac 1920  
DB 1861 GCCGCGAACAGCTGCACACTGACAGCAACAGAGAAACAAATCAAGTCCAGGCCGGCAC 1920  
OY 1921 agccagtgagttcccttcttcccttcttctggaagttgttgaacacccggtlcaagcctct 1980  
DB 1921 AGCCAGTGAGTTCCCTTCTTCCCTTCTTCTGGAAGTTGTTGAACACCCGGTACAGCTCTCT 1980  
OY 1981 tgcattggaacttggaacaattttgtacacgattgtgcgcgtgcctcgaagtaaccccgag 2040  
DB 1981 TGCTATGGGAATTGTGACAAATTTGTACACGATGTGCGCGTCCCTCAAGTACCCCGAG 2040  
OY 2041 cccaaccttggctcccgagcgctgtgtcgtcacaacttataatgatacctgtgtggagc 2100  
DB 2041 CCCAACCTTGGTCCCGAGCGCGGTGTGCTGCATATTACATGAGATCTGTTGGGGC 2100  
OY 2101 agccctcatcctgttcttactctcactcactcactggtgttgaacatcgtctggaagag 2160  
DB 2101 AGCCCTCATCCTGTCTGTACTCTCCTCCTGACTGACCTGCTTGACATCTGTGGAAAGAG 2160  
OY 2161 caagtcctccccaacccccaacagctgcaactgacaaagcagaagagagcagcaagctgc 2220  
DB 2161 CAAGTCTCTCCCAACCCCAACAGCTGACACTGACACAGCAGAGAGGAGGCCAGACTCTG 2220  
OY 2221 caaccaagactgtgacagcaacccaacccaacacgctccttgcaactccctcaactggag 2280  
DB 2221 CACCAACAGACTGTGACAGCAACCCCAACACCCCTCTCTTGCACCTCCCTCCTCAGTGGG 2280  
OY 2281 tggcagcaacagcagctggaactcctcctcactcaacacagagcagatggtacacatgggg 2340  
DB 2281 TGGCAGCAACAGCACTGAGCTCTCTCTTCAACTGACAGCCCAATGTCACATAGGGG 2340  
OY 2341 tgaactgtctcttcttcaatttatttcttctacgaagaagacaggtgcaacccacaga 2400  
DB 2341 TGACATGCTGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2400  
OY 2401 cccctctgagaaccccgctgtgcgcgcaagcagcagccctgttctcctagcagcagagt 2460  
DB 2401 CCTCTGAGAAACCCGCGTGGCGCCCAAGCCAGAGCCCTCTCTAGGCCAGAGGT 2460  
OY 2461 tctagtgaggggtgagcccttcaagccttcaagagtgagtgagcagagccctcccaacaaag 2520  
DB 2461 TCTAGGTAGGGGTGAGCCCTTGTCAAGCCTTCAAGATGGGACAGCCCTCCCAACAAAG 2520  
OY 2521 gtcaacctcaactgaaatgtlacaacacccaacagctgtlccaaaggccttagtccctactt 2580  
DB 2521 GTTCAACCTCAAACTGAAATGTAAACAAACCAACCAAGCTGTGCAAAAGGCTAGTCCCTA 2580  
OY 2581 ctgctactgtctgtcctgtgagcctctgaagggcccccctcaataaaaagcttgaacagcag 2640  
DB 2581 CTGCTACTGTCTGTCTGTGAGGCCCTGAAGGCCCCCTCATCAAAAAGCTTGAACAGGCGAG 2640  
OY 2641 cccagagtggtlcaacccctgggctactgggaggaagaaacaaacactcaaaagtctgcaac 2700  
DB 2641 CCCAGAGTGTTCACCCCTGGGCTACTGGGGCAGAAAGAAACCTTAAGATGTGCACAC 2700  
OY 2701 acaacaaggaagcgtlccctcctcctgataagctgacataagcctgtgtgtlctgctacaat 2760  
DB 2701 ACACACAAGGAAGGCTCTCTCTCGAATAGCGAATAGGCGCTGTGTGTGCTTACACT 2760  
OY 2761 tcatgttctacttaatcctctcaagaagcaacccctgggaagagcctcgcaggaacctc 2820  
DB 2761 TCATGTTCTACTTAATCTCTTCAAGACAGCAACCTCGGGAAGAGACCTCGCAGGACCTC 2820  
OY 2821 cccagacaagaagaagaagaagaagaagggtgatatataagcagcgagttccct 2880  
DB 2821 CCCAGACAAGAAAGAAAGCAAGAGGAGGATTAATTAAGCAACAGCAATTTTCCCT 2880  
OY 2881 atcccttaacctagagtcaccacactgaaatgccaacagcctgccaagaaaccttgac 2940

```

|||||
Db 2881 ATCCCTTACCTAGAGTCCCACTGAATGACACAGCCTGCACAGAAACCCCTGGC 2940
QY 2941 aaagcttgagctgctctgtgcaacccctctgacctgctcagggaatacagggccctcag 3000
Db 2941 AAAGGCTGGAGCTGCTCTGTGCCACCTCCTGACCTGTGAGGGAATCAGAGGCCCTCAG 3000
QY 3001 gcaagctggagacagagctcctctgtcctcagtaatactcctgtctgagtgccctc 3060
Db 3001 GCAGCTGGAGACAGGCTCTCTCTGTGCCATCAGTAATCTCTGTGCTGGATGGCCCTC 3060
QY 3061 cccacacctataaattctctgagacaccttgctgcaagaataaagtgcttctt 3120
Db 3061 CCCACCTTATATTAATCTCTGAGTACACCTTGCATAGAAATAAAGTGTTGCTTT 3120
QY 3121 gtaa 3124
Db 3121 GTAA 3124

RESULT 2
BC003964 2336 bp mRNA ROD 16-MAR-2001
DEFINITION Mus musculus, clone IMAGE:3488182, mRNA, partial cds.
ACCESSION BC003964
VERSION BC003964.1 GI:13278270
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2336)
REFERENCE Strausberg, R.
TITLE Direct Submision
JOURNAL Submitted (28-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Lotmar Hemighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) medepaxj.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAK Plate: 7 Row: k Column: 18.
Location/Qualifiers
1. 2336
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3488182"
/tissue.type="Mammary tumor. MAP-TGF alpha model. 7 months
old, gross tissue."
/clone_id="NC1_CGAP_Mam5"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
<1. 1048
/codon_start=2
/product="Unknown (protein for IMAGE:3488182)"
/protein_id="AAH03964.1"
/db_xref="GI:13278271"
/translation="GYMLFERNALLESYLAHKKYIKPSGNMPTIGDVHLAPFTDEOLY
MEQFTANFWYQPSFGVDLSALRGAAVDEYFQPVVDIDILMAKSVKYVNFLE
AKRGDLHRIEIPKFMHLHSGLVHGLAFMVDVAFISIMVWLSTAPTEBLTHMYOVR

```

```

CLFQSPLEFAKAGDTLSGTCILLIANKROSDYISIVAYQDGTGSKSNLDJKNPFERYT
GTPSPSPGSHYTSPEENMNTGTSYLSGVAVAGAPTAADYLSVYAGSGSVGHNNL
IPLANGIVNHTHSRMSIMSTGTYGSSSAQGGGSSSAHYANNNOFTMGPAISMA
SPMSIPRNTMHYGS"
BASE COUNT 563 a 740 c 537 g 496 t
ORIGIN
Query Match 72.2% Score 2256.4; DB 94; Length 2336;
Best Local Similarity 99.3% Pred. No. 0;
Matches 2308; Conservative 0; Mismatches 11; Indels 5; Gaps 4;
QY 804 gggctacatgctcttcacatgacgaatgctcagagctacctccatgccaagaatct 863
Db 1 GGGCTACATGCTCTTCAATGATGACGAATGCTGAGAGCTACCTCCATGCCAAATATCT 60
QY 864 gaagctagtggaacatgctcccaacatgtgtaatgctacacctcgacacctcaatga 923
Db 61 GAAGCTAGTGGAACATGTTCCCAACCATTTGGATGTCCACCTCCGACCTTCATGTA 120
QY 924 tgaacagctctacatgagcaggttcacccaagccaactccggtaccagccatcttcca 983
Db 121 TGAACAGCTCTACATGAGCAGTTCACCAAGCCAACTTGTGTACCAAGCCATCTTCCA 180
QY 984 tggagtggaacctgctgcgcccacagaggtgcccgtgtgtagtacttcggcaacctgt 1043
Db 181 TGGAGTGGACCTGTGCGCCCTCAGAGGCGCGGTGTGATAGTACTTCCGGCAACTGT 240
QY 1044 ggtgacacatttgacatccgatactgataagccaatctgtcaagtaacagatgaact 1103
Db 241 GGTGACACATTTGACATCCGGATCCTGATGCGCAAACTGTGCAAGTACAGTACACTT 300
QY 1104 cttagaagccaagaagcgatgtgcacagatagaatacccatcaaatccacatgct 1163
Db 301 CTTAGAAGCCCAAGAAGCGATTTGACAGATGAAATCCCATCAATTCACATGCT 360
QY 1164 gcatcagaggtagtcacatgctgcttgccttgcgtgtgtagtattgttccattgctcat 1223
Db 361 GCATTCAGGGCTAGTCCATGGCTTGCCCTTGCTTGCTGATGTTGCTTTCATTGGCTCAT 420
QY 1224 aatgacgtgtgctatccacagccccaagagagcccttgaccccaactgtgtaacagctcg 1283
Db 421 AATGACGTGTGGCTATGCCACAGCCCAAGAGCCCTGAGCCACTGTGATCAGAGTCCG 480
QY 1284 gtgcctcttcacgaacccgtgtgttgcgaagcgcggaacagctctcaaggatgctc 1343
Db 481 GTGCCTTTCAGTACACCGTGTGTTGCCAAGGCCGGGACACGCTCTCAGGACATGTCT 540
QY 1344 gctattggccaagaagacagagctatgacatagatattgtgacacaggtgagcaagac 1403
Db 541 GCTTATTGCCAACAAMACACAGACTATGACATCATGTAATGTGGACACAGGTGACACAGAC 600
QY 1404 aggtcccaagtcagtaacccgtgtgataaagaacccctctcagaatgacagagttac 1463
Db 601 AGGCTCAAGTCCAGTACACCGTGTGATCTTAAGAACCCTTCTCAGGTACACAGTAC 660
QY 1464 aaccocatcaccccaacctgtgctacactacagctccctctcgagaatatgtgaaacac 1523
Db 661 AACCCCATCACCCCACTGCTCAGACTACACGCTCTCCCTCGGAGATATGTGGAACAC 720
QY 1524 aggaagacacataatctcaacagcggtgtgctgtgctggaatgactacgctcactaga 1583
Db 721 AGGAAGCACTATATCTCAGACAGCGGGGTGTGTGCTGGAATGCTTACTGCTTACGA 780
QY 1584 cctgagcagtttatctgcgagcggtccagtggtgtgtgtacaaacacactgattcccttagc 1643
Db 781 CCTGAGCAGTGTATTTGCGGGGGGTCCAGTGTGGGTACAAACACACGATTCCTTAGC 840
QY 1644 taacacagaggtatgcaatacaacacccactcccgatggtgtccataatgagcagggcat 1703
Db 841 TAAACAGGGGATTTGCAATACACCCACTCCCGGATGGGCTCCAAATAAGACACGGGAT 900
QY 1704 tgtcgaagctccctcaggtgtgccagggagcggtgtagtcacagtgcccaactagcagt 1763

```

|    |      |                                                                    |       |
|----|------|--------------------------------------------------------------------|-------|
| Db | 901  | TTGTCCAAAGGCTCTCTACAGTGTGCCAGGGAGGGCGGGAGTCCAGTCCAGTCCACTATGAGT    | 960   |
| Oy | 1764 | caaaacaaccggttcacaaatgggtgtgcccctgcatactctatgtgcctcgcccatgtccatccc | 18233 |
| Db | 961  | CAACAACCGATTACACATGGGTGGCCCTGGCCAACTCTATGGCCCTGGCCAAATGTCCATACC    | 10200 |
| Oy | 1824 | gaccaaaccatgtcaactatgggaggttagtgtgtcctccagccgagcaagacatgtcgcattga  | 18833 |
| Db | 1021 | GACCAACACCAATGCACCTATATGGAGTTATAGTGTCTCCAGCCGGCAGACACACTGGCCACTGA  | 10800 |
| Oy | 1884 | cagcaaccgagaaacaacaatcaaatgccagggccggcagcagcagatgtgtgtgtcccccgtgt  | 19433 |
| Db | 1081 | CAGCACACAGGAACCAAAATCAAGTCCAGGGCCGGCAGCCAGTGGGTGTCCCCCTTGT         | 11400 |
| Oy | 1944 | tctgagaagaagtgtgtgaacaaccggttcaag - cctcctgtctatgggaacttggacaat    | 20022 |
| Db | 1141 | TCTGGAGAAGT - GTGAAACACCCGGGTACAGCCCTCTTTTGCTATGGGAACATGTGACAAAT   | 11988 |
| Oy | 2003 | ttgtacacagatgtgcgcgtgtccctcaatgaatcccccaagcccaacotttgtlcccgagcgc   | 20622 |
| Db | 1139 | TTGTACACCATGTGCGCGGTGCCCTTAAGTACCCCAAGCCCAACTTGTGGTCCGACGCG        | 12588 |
| Oy | 2063 | gtgtgtgtgcatacttcaatcatgatactcctgtttgggcaagccctcaatcctgttctgact    | 21222 |
| Db | 1259 | GGTGTGCTGCCATCTTATCATAGATCCGTGTGGGGACACCCCTCAATCCTGTCTGACT         | 13188 |
| Oy | 2123 | ctccacctgcacctgtgctttgacatctgtctggaagagcaagtcctccccaaccccac        | 21822 |
| Db | 1319 | CTCCACTCTGACCTGGCTTGTGACATCTGCGTGAAGAGGCAAGTCCTCCCAACCCCCAC        | 13788 |
| Oy | 2183 | agctcacctctgaccagagagggagagcagcagctgtccacacacagacctgtcagacacc      | 22422 |
| Db | 1379 | ACTGTGACCTGTGACACAGGCAAGAGAGGCAAGCAGCTGCCACACAGCCTGGCAGCACCC       | 14388 |
| Oy | 2243 | accccaacaaccgcttcccttgacacctccctcaacctgtgggtgtgcagcaagcagctgac     | 23022 |
| Db | 1439 | ACCCCAACACCCGTCCTTTCACACTCCCTCCTACCTGGGGGTGGACACAGCAGCTGGAC        | 14988 |
| Oy | 2303 | tctccttcaactacagagccaaatgttcaacatltgggcgtgaacatgtcgtctttttaat      | 23622 |
| Db | 1499 | TCTCCTTCAACTACACAGGCGACATGTGTCAACATGAGGCGTACATGCTTTTATAT           | 15588 |
| Oy | 2363 | ttaattttttagaanaagaaacagatgtcaaaccaagaccctcttgaaaaacccggcgcg -     | 24212 |
| Db | 1559 | TTATTTTATTTAGAAAAGAACCACTGTCAACCCACAGACCCCTTGAAAAACCCGGCGTGC       | 16188 |
| Oy | 2422 | cgcgccaagccagcagcccccgtctccctagcccaaggttctatgttgaagggtgtgcctgt     | 24812 |
| Db | 1619 | CGGGCAACCCAGCAGACGCCCTGTCTTAGGCCAAGAGTTCTTAGTGAAGGGGTGGCCCTG       | 16788 |
| Oy | 2482 | tcaagccttcagagtgtggcacag - ccccttccacaacaaggttcaacctcaacttgaatg    | 25402 |
| Db | 1679 | TCAACGCTTCAAGATGGGCAACAGCCCCCTCCACCAAAAGGGTTCACCTCAAACTTGAATG      | 17388 |
| Oy | 2541 | tacaaacaccccaagctgttccaaagcctagtccctacttctgtctactgttccgttctga      | 26002 |
| Db | 1739 | TACAAACCAACCCAGCTGTCCAAAGGCGCTTAGTCCCTACTTCTCTACTGTGCTCTGA         | 17988 |
| Oy | 2601 | gccctgaaagcccccttccatacaaaagctgtgaacagcagcccaagatgtgtgaacccctgtg   | 26602 |
| Db | 1799 | GGCCTGAAAGGCCCTCCATCAAAAAGCTTGAACAGGACCCCAAGATGTGTCAACCTGG         | 18588 |
| Oy | 2661 | gctactgtggcagaacaagaacctccaagaatctgttcaacacacaacaagaagtcgtcct      | 27202 |
| Db | 1859 | GCTACTGGGGCAGACAAGAAACTCTAAAGTCTGTCAACACACACAAGAAAGCGTCTCT         | 19188 |
| Oy | 2721 | ctcctgtatgtctgaatagcctgtgtgtgtgtgttcaacatlaatgttctaataatcc         | 27802 |
| Db | 1919 | CTCCGTATGTCTAATCATAGGCCGTGTGTGTGCTTCAATTCATGTCTACTTAATCTCTC        | 19788 |
| Oy | 2781 | tcaagaacagcaacccctgggaaaggacctcgcaaggaccttcccagacaaagaagaaaaagca   | 28402 |

```

Db 1979 TCAGAGCACCAACCCCTGGGAAGAGAGCTGACAGGAGACTCCACAGCAAGAAGAAAAAGCA 2038
Qy 2841 aacaaggaaagggtgatataatgaacacaggcaagttccctattcccttaccctagagtc 2900
Db 2039 AAGAAGGAAGGGGTGATTATTAAGCACAGGCAGATTCCCTATTCCCTTACCCTAGAGATCC 2098
Qy 2901 ccacctgaattgccaagaagctgccacaggaaacctctggcaaaagctggaagctctgt 2960
Db 2099 CCACCTCAATGGCCACACAGCTGTGCACAGAAACCCCTTTGGCAAAAGGCTGGAGCTGCTGT 2158
Qy 2961 gccacctctctgacctgctcaaggaaatcaaaagggccctcagcagctggaaccaagctct 3020
Db 2159 GCCACCCCTCTGCTGACCTGTCAGGGAATCAGAGGGCCCTTCAGAGCAGCTGGGAACAGGCTCT 2218
Qy 3021 ctctcgtcatcaatgaatactctctgctcgatgagcctcccccaccttatataatc 3080
Db 2219 CTCGCTGTCATGAGAAATACCTCTTGCTGTGGCGCCCTCCACACTTATATATAAATTC 2278
Qy 3081 tctgatacccttgcataagaaaataaagtgtttcttgtaa 3124
Db 2279 TGTGATCACCTTTTGCAATGAAAATAAAGTGTTCCTTTGTA 2322

RESULT 3
AC073731
LOCUS AC073731 233830 bp DNA HTG 29-JUN-2000
DEFINITION Mus musculus clone RP23-247F20, WORKING DRAFT SEQUENCE, 20
unordered pieces.
AC073731
AC073731.1 GI:8810348
HTG: HTGS.PHASE1: HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS 1 (bases 1 to 233830).
TITLE DOE Joint Genome Institute.
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 233830)
TITLE DOE Joint Genome Institute.
JOURNAL Direct Submission
AUTHORS Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT

Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1831860
Center clone name: RPC1-23_247F20

Summary Statistics
Consensus quality: 215926 bases at least Q40
Consensus quality: 224967 bases at least Q30
Consensus quality: 226764 bases at least Q20
Estimated insert size: 225370; agarose-fp estimation
Estimated insert size: 231930; sum-of-contigs estimation
Quality coverage: 7.83 in Q20 bases; agarose-fp estimation
Quality coverage: 7.61 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1072: contig of 1072 bp in length
* 1073 1172: gap of unknown length
* 1173 2301: contig of 1129 bp in length
* 2302 2401: gap of unknown length

```



```

Db 96032 TTTGTRA 96038
|||||
RESULT 4
AF055027 2160 bp mRNA PRI 02-APR-1998
LOCUS Homo sapiens clone 24658 mRNA sequence.
DEFINITION AF055027
ACCESSION AF055027.1 GI:3005756
VERSION
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2160)
Anderson, B., Wentland, M.A., Rierafrante, J.Y., Liu, W. and Gibbs, R.A.
TITLE A 'double adaptor' method for improved shotgun library construction
JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
MEDLINE 96207227
REFERENCE 2 (bases 1 to 2160)
Yu, W., Anderson, B., Morley, K.C., Muzny, D.M., Ding, Y., Liu, W.,
AUTHORS Rierafrante, J.Y., Wentland, M.A., Lennon, G. and Gibbs, R.A.
TITLE Large-scale concatenation cDNA sequencing
JOURNAL Genome Res. 7 (4), 353-358 (1997)
MEDLINE 97264341
REFERENCE 3 (bases 1 to 2160)
Yu, W. and Gibbs, R.A.
AUTHORS Direct Submission
JOURNAL Submitted (23-MAR-1998) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
FEATURES
source
1. 2160
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="brain"
/dev_stage="infant"
/sex="female"
/clone_id="11NB"
/clone_image="IMAGE Consortium 24658"
BASE COUNT 468 a 717 c 551 g 424 t
ORIGIN
Query Match 36.5%; Score 1141.2; DB 88; Length 2160;
Best Local Similarity 81.2%; Pred. No. 1.2e-235;
Matches 1504; Conservative 0; Mismatches 263; Indels 85; Gaps 12;
QY 727 cgcacgtggtcatccctgycacaaagtagagaggtctcatctgctgagcaagtggacatt 786
Db 1 CGCATCGTGGTCATCCCGGGCAAGTGGAGAGGTGTCACTCCCGCAGACAGTGGAGATC 60
QY 787 atcattcagagccatgggtgtacatgctctctcaatgaaagatgctcagagatactc 846
Db 61 ATCATCTCGAGCCCATGGGTACATGCTCTTCAACGACGCGATGCTGAGACTACCTC 120
QY 847 catgcaaaaagttacctgaagcttagtgaacaatggttcccaaccttggtgattgcac 906
Db 121 CACGCCAAGAACTACTGAAGCCCAAGCGAAGCAATGTTCTTCAACATTGGTGCCTCAC 180
QY 907 ctgcacaccttaactgataagcagctctatcagtgagagcagttaccacaagcaactccag 966
Db 181 CTTCGACCCCTTACAGGATGAGACAGCTCTACATGAGACAGTTCACCAAGGCCAATCTTCG 240
QY 967 taccagcatccttcacatggagtgagacgtgctgcgcctcagaagtgctcgctgtgagtag 1026
Db 241 TACCGACCATCTTTCATAGAGAGTGAACCTGTGCGCCCTCCGAGGTGCCCGGGGATGAG 300
QY 1027 tacttcggcaacctgtgtgtgacacatttgacatccggaatcctgattgagcaaatctgac 1086
Db 301 TATTTCGGGCAACCTGTGTGTGACACATTGACATCCGATCTCTGATGCGCAAGCTGTGC 360
QY 1087 aagtaacagtgaactctttagaagcaagaagcagatttgcaagagtagaataccca 1146

```

```

Db 361 AAGTACAGGCTGAACCTCTTAGAAGCCAAAGACGATTTGACAGGATGAATCCCA 420
QY 1147 ttcaaatccacatgctgcattcaggttagtccatggtcttgccctctgttgcagt 1206
Db 421 TTCAAATTCACATGCTGATTCAGGGCTGTCTCAGCGCTGGCTTCTGTTGACCTT 480
QY 1207 gcttcattgctccacaatgacccgtgtggtatccacaagcccaagcccaagccctgac 1266
Db 481 GCTTTCATCGGCTCCATTAATGACCGTGTGCTGTCCACAGCCCGGACAGACCCCTGAC 540
QY 1267 cactgtgaccaggtccggtgctctctccagtcacagctgtgttgcagaagccggagacg 1326
Db 541 CACTGTGATCAGAGTGGCGGTCTGCTCTGTTCCAGTACCACTGTTCGCAAGGACAGGAGACG 600
QY 1327 ctctcaggagacatgctgcttatttgcccaaaaaagagagctatgacatgaattgtg 1386
Db 601 CTCTCAGGACATGTCTGTTATTGCGCAAAAGACAGACTGACATGATATTGTG 660
QY 1387 gcaacagtgaccagagcgtcccaagttccagtaacctgtgagttctaaagacccttc 1446
Db 661 GCCCAGGTGACACAGACCGGCTCCAACTCCAGTAACTCTTGATTTGAATAACCCCTTC 720
QY 1447 ttcagytacacaggtacaaacccatcaccccaactgtgctacacatacagctctccctcg 1506
Db 721 TTTAGATACAGGGGACACAGCCCTCACCCGCCGCTCCACTACATCTCCCTCG 780
QY 1507 gagaatatgtggaacacaggaagacactataatcagaacaggggtgtgctgtgtgga 1566
Db 781 GAAACACTGTGGAAACAGGGGACAGCACTTACACCTCAGCAGCGGATGGCGTGGCAGGG 840
QY 1567 atgcctactgctcctcagacccctgagcaggtgtattgcccggcgctcagttgtgtgtgacac 1626
Db 841 ATGCGACCGCCCTTATGACTTGAAGCAGTGTATTGCCAGTGGCTCCAGCGTGGCCACAC 900
QY 1627 aacctgattcctttagtgaacacagggatgtgtcaatcacaccactcccgatgtgctcc 1686
Db 901 AACCTGATCTCTTATGACCAACGGGATGTTGTCAATACACCCACTCCCGGATGGGCTCC 960
QY 1687 ataatgacagggagattgtccaaaggtcctcaggtgccaaag---gagcgcggtgac 1743
Db 961 ATATGACGACGGGGATTTGTCCAAAGGCTCTCCGGCCCGCAGGGCAGTGTGTGGCGAC 1020
QY 1744 tccagtgcccatgatagaatgacaacaaacagttcacatggtgtggtgctgcatctatg 1803
Db 1021 ACAGATGCCCCATATGACATGACATACAGCCAGTTCCATGCGGCGCCCGCATCTCATG 1080
QY 1804 gcttcgcccattgtccatcccgacaacacacatgacatggaattgagttgctcccaagcc 1863
Db 1081 GCGTGGCCCATGTGCCATCCCGACCAACCATGACATGACGGGACTGAGGGC----- 1132
QY 1864 gcgacagcactgacgacatgacagcaacaggaacaaatcaatgaccagc---ccggcac 1920
Db 1133 ----CCGCCCCGGGAGCTGACACACAGGAAACCAATGATGTCCTGCCGCCGCCGCC 1188
QY 1921 agccagtgctgtgtcccccgtgtgttgagaagtgtgtgaacacccggtacagc-ctcc 1979
Db 1189 CGCGGGCGGCTTCCCTTGTACTGTGAGAG--CTCGAACACCCCGGTACACACTCTCT 1246
QY 1980 ttgctatgggaacttggacaatttggacagatgtgcgcgtgcgcctcgaagtaagcccca 2039
Db 1247 TTCTTATGGGAATGGGACACTTTTTCACGATGTTGCGCGCTCCCA---CCTTAAC 1303
QY 2040 gcccaaccttggctcccgagcggtgtgtgtgcataacttaacatgatacgtctgtggg 2099
Db 1304 CCCACCTCCCGGCGCTGAGCGTGTGCGCTGCACATTTTACACAATAATCATGTTGTGG 1363
QY 2100 cagccctcat--ctgttctgtactcactcactgacactgtgtgacatctgtcgtgaag 2157
Db 1364 GAACCTCTGTCCCGCTCTCTGCGCGCGCTTACCTGTACCTGGGCTTGTTCATCTGTGTGAAC 1423
QY 2158 aggcagagtcctcccccacaccccaagctgacacttgacacagcgacggagggagcgacagc 2217

```





QY 3124 a 3124  
 Db 1114 A 1114

RESULT 6  
 LOCUS BC002282/c  
 DEFINITION Mus musculus, similar to hypothetical protein, clone MGC:7695,  
 mRNA, complete cds.  
 ACCESSION BC002282  
 VERSION BC002282.1 GI:12805604  
 KEYWORDS MGC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2107)  
 Strausberg, R.  
 Direct Submission  
 Submitted (31-JAN-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www.sphg.stanford.edu>  
 Contact: (Dickson, Mark) [modexpi1.stanford.edu](mailto:modexpi1.stanford.edu)  
 Dickson, M., Schultz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

REMARK  
 COMMENT  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 8 Row: 1 Column: 24.  
 Location/Qualifiers  
 1. 2107  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="MGC:7695 IMAGE:3497525"  
 /tissue\_type="Mammary tumor. WAP-TGF alpha model. 7 months  
 old, gross tissue."  
 /clone\_1lb="NCI-CGAP\_Mam5"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 189..1127  
 /codon\_start=1  
 /product="Similar to hypothetical protein"  
 /protein\_id="AAH02282.1"  
 /db\_xref="GI:12805605"  
 /translation="MAADELAFHEFEENATLAEFPDAATTSOSDELSTREHAVVY  
 GSGIGYGAVEEDDKTSLOEKRPQPFMTWDDYOSFDDVTSIDPKIKSLDPH  
 PGHNFRHHLRRRDLXGPKFCATLAPLAYGNTLYLAQRDSIRHSPPHKYT  
 IAGITTYCTAMLYPLALMGFLRMROGTRFMGLYTLLEVVCYVGSLEFVTPVIML  
 IIVQWQLGALGLALSAAGLFTLMPVREDTRVLAALSTVLLHLALMGCKL  
 YFQPLPDHVPAPDAPPPSPVLLPSSIQPTWS"

BASE COUNT 447 a 533 c 608 g 519 t  
 ORIGIN

Query Match 23.9%; Score 745.8; DB 94; Length 2107;  
 Best Local Similarity 98.8%; Pired. No. 1.9e-150;  
 Matches 772; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 2346 tgcgtgttttttaatttttttttaagaaagaaagcagtgtaaccacagaccctc 2405  
 |||||||  
 Db 2092 TGCTGCTTTTAAATTTATTTTAAACGAAAGAACAGTGTCAACCAACAGACCCCTC 2033

QY 2406 tgaagaaccggcgtg-cgcgcgaagccagcagccctgtctcctaagccagaagttcta 2464  
 |||||||  
 Db 2032 TGAGAAACCCGGCTGGCGGGCCAAACGACGAGCCCTGTCTTAGGCCAGAGGTTCTA 1973

QY 2465 ggtgaggggtggccctgtlcaagccttcagagtgggacag-ccctcccaacaaaggtt 2523  
 |||||||  
 Db 1972 GGTGAGGGGTGGCCCTGTCAAGCCTTCAGAGTGGGCACAGCCCTCCACCAAGAGGTT 1913

QY 2524 caccctaaactgaaatgtacaaaccccaagctgtccaaagctgtccctcttctg 2583  
 |||||||  
 Db 1912 CACCTTAAACTTGAATGTACAAACCAACCAAGCTGTCAAAGGCTGTCCCTTCTTG 1853

QY 2584 ctactgcctgcctcgaagcccgaaagcccccctccatcaaaagcttgaaagcaagcc 2643  
 |||||||  
 Db 1852 CTACGTCTCTGCTCGAGAGCCCTGGAAGGCCCTCCATCAAAAGCTTGAACAGGCGCC 1793

QY 2644 agagtgtgtcaccctgggctactggggcagacaagaaccccaagatctgtcacaca 2703  
 |||||||  
 Db 1792 AGAGTGTGCACCTGTGGCTACTGGGGCAGACAAAGAACCTCAAGATCTGTACACACA 1733

QY 2704 cacaaggaaggggtcctcctcctcgaatagctgacataagcctgtgtgtgcgttcaatca 2763  
 |||||||  
 Db 1732 CACAAGGAAGGGCTCTCTCTGTGATAGTAAACATAGGCTGTGTGTTGGTTCAATTC 1673

QY 2764 tgttacttaactctctcaagaacagacacccctgggaaggaagcctgcagagaccctccc 2823  
 |||||||  
 Db 1672 TGTCTACTTAATCTCTCTCAACAGCAACCTGGGAAGAGGCTGTGAGGAGCTCTCC 1613

QY 2824 agacaagaagaagaacaaacaaaggaaggtgtatlaaagcagaagcagttccctatt 2883  
 |||||||  
 Db 1612 AGACAAGAAGAAGCAAGAAAGGAAGGTGATTAAGACAGGCAATTTCCCTATT 1553

QY 2884 ccctaacctcgaagctcccaactgaatggcacaagcctgcacagaacccctggcaaa 2943  
 |||||||  
 Db 1552 CCTTACCTTAGAGTCCCACTCAATGCGCACAGCCTTCCACAGAACCCCTTGACAA 1493

QY 2944 gctggaagctcgtcgtgcacacccctcgaactgtcaaggaatcaagggccctcaagga 3003  
 |||||||  
 Db 1492 GGCTGAGACTGCTGTGTGACACCTCTGACCTGTGACAGGAATCAGAGGCCCTCAGGCA 1433

QY 3004 gctgggaacagagctctctcctgtccatcaagtaactcctgtcgtcgtgaatggccctccc 3063  
 |||||||  
 Db 1432 GCTGGGAACAGAGCTCTCTCTGCTCAGTCAATCAATATCTCTTGTGCTGCCCTCCCC 1373

QY 3064 cacccttatataactctcctggaatcaacttgcataagaataaagtttgccttga 3123  
 |||||||  
 Db 1372 CACCTTTATATTAATTCCTGTGATCACCCTTGCAATAGAAATAAAGTGTGCTTGTGA 1313

QY 3124 a 3124  
 Db 1312 A 1312

RESULT 7  
 LOCUS AX079470  
 DEFINITION Sequence 214 from Patent WO0107611.  
 ACCESSION AX079470  
 VERSION AX079470.1 GI:13159030  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 500)  
 Baker, K.P., Goddard, A. and Wood, W.I.  
 Human polypeptides and methods for the use thereof  
 Patent: WO 0107611-A 214 01-FEB-2001;  
 Genentech, Inc. (US)  
 Location/Qualifiers  
 1..500  
 /organism="Homo sapiens"











```
QY 929 agcttcatatgagcagcttaccacaaagccactcccggtaccagccatcccttccatgag 988
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48253 CGCTCTACGGGAACAAATATAACAAGGCCAACTCTGTGATCAGTCGGCTTCACAGCGG 48194
QY 989 tggacactgtgagcctcagaggtgagcgtgtgtgagagacttccggacactgtgtgtg 1048
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48193 TTGACCTTACCAACGCTGCACAAAGAGGCGATAGAGACTTCCGACGCCCATTTGTGG 48134
QY 1049 acacattgacatcccgatccctgcatgacaaatctgtcagtaacagagaaactcttg 1108
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48133 ACACCTTGACATCCGCAATATCATGACCAACTCGGTGGGACAGCGTCCGACTTCTCA 48074
QY 1109 aagccaaagaggcgatltgacagagat----- 1136
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48073 ACGATTAAGAGAGAGCATGCTGATTGTGAATAGTACCCGAAAAAGTATTGCGACAA 48014
QY 1137 -----agaaatcccatlcaaatlccaatlccgcatlccagc 1174
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48013 TGGCCAACTGTTGACTTCCGACGACATCCCTCGAGTTTCAATCTGCAAACTGGCA 47954
QY 1175 tagtccatgctgtgctccttctgtcgtatgtgtcttccatctgtccataatgacctgt 1234
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 47953 TCTGCGAATGGGCTGGCTTTTGTTGATGTGAGATTAGCGGACAGCAGCAATGCT 47894
QY 1235 ggcatacacagccccaaacagagcccttgacccactgtgtacaggtccggttcccttcc 1294
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 47893 GGCCTCAACGTCGCCACAGCGCCTCTGACACATTTGGTATCAGGTTCCGCTCTCTGC 47834
QY 1295 agtaacgtgtgttgacagagccggagccgtctccagggacatgtcgtctattgtgca 1354
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 47833 CCATGCCATCTTTATTTAAGCAGGGGCCAGACGCTGACCGCGGTCTTCTGAGGAGCCA 47774
QY 1355 acaaaagacagatcatatcatatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1414
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 47773 ATGGCGGCGCATGCTATATGACGTGACCATGACCTACATCAAGGACAGCTGATTTGCT 47714
QY 1415 ccaagtaacctgtgtatctaaagaaccccttctcaggtacacaggtacacacccatcac 1474
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 47713 CGAGCAACACCTGAGACCTTTAAGAAATCCGTATTTTCGTACACGGGGGCCCGGTCACAG 47654
QY 1475 cccacacgt 1532
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 47653 CACCGCGGGGCTGTGATCCCAAGTCCATCTGAGCATGAGACACAAAGTGAGACAC 47586

RESULT 11
AX079494 490 bp DNA PAT 22-FEB-2001
LOCUS AX079494
DEFINITION Sequence 238 from Patent WO0107611.
ACCESSION AX079494
VERSION AX079494.1 GI:13159054
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 490)
AUTHORS Baker K.P., Goddard A. and Wood W.I.
TITLES Human polypeptides and methods for the use thereof
JOURNAL Patent: WO 0107611-A 238 01-FEB-2001;
Genentech, Inc. (US)
FEATURES
Source
1. .490
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 111 a 150 c 124 g 92 t 13 others
ORIGIN
Query Match 10.5%; Score 329.4; DB 10; Length 490;
Best Local Similarity 85.3%; Pred. No. 1.4e-60;
Matches 371; Conservative 0; Mismatches 63; Indels 1; Gaps 1;
```

```
QY 1219 tccataatgacggt 1277
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TCCATATATGACCGCTGTGAGTGTCCACAGCCCCACAGAGCCCCCTGTGACCAATGTACCA 60
QY 1278 ggtcgcgtgctccttccatgacacacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1337
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GGTGCGGCTGTGTTCATGACCATTTGTGCGCAAGGAGGAGGACAGGNTTCAAGGAGAC 120
QY 1338 atgtcgtctattgt 1397
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ATTGTTGNTATTGCGACAAAGACAGAGNTAGCATATGATGTGGCCAGGAGGGA 180
QY 1398 ccagacagctcccaagttccagtaacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1457
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CCAGACCGGCTCCACATCCGATTAACCTCCGATGATGAAAAACCCCTTNTTGTATACAC 240
QY 1458 agttacacccac 1517
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GGGCACACAGCCCTCACACCCACCGGNTCCANTACATNTCCCTCGAAAAACATGTG 300
QY 1518 gaacacaggaagacacataatctcagcagcgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1577
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GAACACAGGGGAGACCTACACACTCAGACAGGGGATGCCGTGACAGGATTCGACGCGC 360
QY 1578 ctacgacactgagcaggt 1637
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 CTTTGTGCTTGTGAGCATGTATTTATTCAGTGTGCTCCAGCGTGCGCACAAACCTGATTC 420
QY 1638 cttagtaacacag 1652
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 TTTAGGCTCTCCGG 435

RESULT 12
AC011442 14357 bp DNA PRI 26-OCT-2000
LOCUS AC011442
DEFINITION Homo sapiens chromosome 19 clone CTC-21504, complete sequence.
ACCESSION AC011442
VERSION AC011442.5 GI:11024828
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 14357)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLES Direct Submision
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 14357)
AUTHORS DOE Joint Genome Institute.
TITLES Direct Submision
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 14357)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLES Direct Submision
JOURNAL Submitted (26-OCT-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Oct 26, 2000 this sequence version replaced gi:9256292.
Draft Sequence produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.1.
STS Content:
SHGC-100746 G55413.
Location/Qualifiers
1. .14357
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTC-21504"
```



```

repeat_region 4119..4419 /rpt_family="Alub"
repeat_region 4844..5142 /rpt_family="Aluy"
repeat_region 5163..5437 /rpt_family="AluJo"
repeat_region complement(5441..5611) /rpt_family="Aluy"
repeat_region complement(5612..5636) /rpt_family="AT_rich"
repeat_region 5716..6059 /rpt_family="L1"
repeat_region complement(6107..6260) /rpt_family="LINE2"
repeat_region complement(6617..6914) /rpt_family="Alusp"
repeat_region complement(6922..6997) /rpt_family="MIR"
repeat_region 7157..7456 /rpt_family="AluSc"
repeat_region complement(8320..8426) /rpt_family="(GGGA)n"
repeat_region complement(8431..8564) /rpt_family="FLAM_C"
repeat_region 8568..8694 /rpt_family="LINE2"
repeat_region 8696..8954 /rpt_family="Alusg/X"
repeat_region complement(8955..8976) /rpt_family="AT_rich"
repeat_region 9028..9324 /rpt_family="Alusx"
misc_feature 9505..9813 /note="predicted exon, program: grr12exons_human_1.3, frame: 2, quality: excellent, score: 81.000"
misc_feature 9551..9637 /note="DPS similarity to AA461913 v95h04.r1 Soares mouse mammary gland BDMG Mus musculus cDNA clone 851575 5'; (1..87): 82% identity."
misc_feature 9727..9813 /note="DPS similarity to AA461913 v95h04.r1 Soares mouse mammary gland BDMG Mus musculus cDNA clone 851575 5'; (88..175): 80% identity."
repeat_region complement(10139..10223) /rpt_family="MIR"
repeat_region complement(10300..10596) /rpt_family="Alusg"
repeat_region complement(10597..10797) /rpt_family="AluJo"
repeat_region 10798..10932 /rpt_family="Alusx"
repeat_region 10933..11240 /rpt_family="AluyB8"
repeat_region 11241..11411 /rpt_family="Alusx"
repeat_region complement(11412..11509) /rpt_family="Alus"
repeat_region 11713..11750 /rpt_family="tRNA-gln-CAG"
repeat_region complement(11752..12062) /rpt_family="LINE2"
repeat_region 12175..12307 /rpt_family="AluJo/FLAM"
repeat_region complement(12309..12524) /rpt_family="AluJo"
repeat_region complement(12525..12595) /rpt_family="Alusg/X"
repeat_region complement(12596..12830) /rpt_family="Alusx"
repeat_region complement(12840..13140) /rpt_family="Aluy"
repeat_region complement(13146..13169) /rpt_family="AT_rich"

```

```

repeat_region 13190..13304
/rpt_family="AluNb"
complement(13307..13609)
/rpt_family="AluSx"
complement(14282..14578)
/rpt_family="AluSx"
complement(14582..14755)
/rpt_family="AluSx"
complement(14775..15073)
/rpt_family="AluY"
complement(15074..15209)
/rpt_family="AluSx"
15511..15664
/note="predicted exon, program: grail2exons_human.1.3,
frame: 2, quality: excellent, score: 92.000"
15511..15603
/note="VDS similarity to AA461913 vF95h04.r1 Soares mouse
mammary gland NbmMG Mus musculus CDNA clone 851575 5';
(176..264); 73% identity."
complement(15748..16049)
/rpt_family="AluSx"
complement(16056..16341)
/rpt_family="AluYo"
complement(16487..16612)
/rpt_family="(CARA)n"
complement(16689..16875)
/rpt_family="AluSc"
16876..17146
/rpt_family="AluSg"
17147..17181
/rpt_family="(GAA)n"
17196..17493
/rpt_family="AluSx"
17562..17589
/note="VDS similarity to AA461913 vF95h04.r1 Soares mouse
mammary gland NbmMG Mus musculus CDNA clone 851575 5';
(265..292); 75% identity."
complement(17721..17889)
/rpt_family="MIR"
complement(18183..18481)
/rpt_family="AluSp"
complement(18489..18625)
/rpt_family="(CARA)n"
complement(18631..18929)
/rpt_family="AluSx"
19161..19204
/note="VDS similarity to AA461913 vF95h04.r1 Soares mouse
mammary gland NbmMG Mus musculus CDNA clone 851575 5';
(293..336); 82% identity."
19498..19800
/rpt_family="AluSg"
19806..20106

Query Match 6.6%; Score 204.8; DB 86; Length 42030;
Best Local Similarity 90.2%; Pred. No. 3.3e-34;
Matches 231; Conservative 0; MisMatches 22; Indels 3; Gaps 1.

Db 3 ggggacctgagcgcgaacctaaagtgaacagcgccagcgagcggtgagccgggtgc 62
|||||
Db 30927 GGGGCGCTGGAGCGCGGATCTAAATGACACCGCGCGCG---CGGCGTGGGCGGCGCC 30983
|||||
Db 63 ggggagcgctgtggtgtggtgcggcgccggcggtgcggcggtgcctctgctacatgctgtgtt 122
|||||
Db 30984 GGGCGGCGCGGGGGTGGCGGTCGCCGGGCGGGCGGGGCGCCCTGCGCTACCGTGTGGGTGTT 31043
|||||
Db 123 cccgggagcgccgcctcctcaatactagcgagcgagcgagcgagatccagcgcaagcggtga 182
|||||
Db 31044 CCGCGGCGCGCCCGCTCCATCCATCAGTGGCGACCGGAACGCGGAGATCCACCGGACGCGGA 31103
|||||
Db 183 ggaagagaggtgtgagccttgaagtgtgcggcgagacagagcgagcgagatgaccttca 242
|||||
Db 31104 GCAGCAGGCGCTGCCTCGAAGTGGCGGCGCGCGCGCGGACCTGCGGCGGATCGGCGCTCTTA 31163
|||||

```









\* \* \* \* \* STN Columbus \* \* \* \* \*

FILE 'HOME' ENTERED AT 09:38:09 ON 21 JUN 2001

=> fil .bec

COST IN U.S. DOLLARS

| SINCE FILE | TOTAL   |
|------------|---------|
| ENTRY      | SESSION |

FULL ESTIMATED COST

|      |      |
|------|------|
| 0.30 | 0.30 |
|------|------|

FILES 'MEDLINE, SCISEARCH, LIFESCI, BIOTECHDS, BIOSIS, EMBASE, HCAPLUS, NTIS, ESBIODBASE, BIOTECHNO, WPIDS' ENTERED AT 09:39:14 ON 21 JUN 2001  
ALL COPYRIGHTS AND RESTRICTIONS APPLY. SEE HELP USAGETERMS FOR DETAILS.

11 FILES IN THE FILE LIST

=> s arginine methyltransferase#

FILE 'MEDLINE'

51900 ARGININE

12238 METHYLTRANSFERASE#

L1 50 ARGININE METHYLTRANSFERASE#  
(ARGININE (W) METHYLTRANSFERASE#)

FILE 'SCISEARCH'

43003 ARGININE

8352 METHYLTRANSFERASE#

L2 55 ARGININE METHYLTRANSFERASE#  
(ARGININE (W) METHYLTRANSFERASE#)

FILE 'LIFESCI'

12502 "ARGININE"

3436 METHYLTRANSFERASE#

L3 28 ARGININE METHYLTRANSFERASE#  
("ARGININE" (W) METHYLTRANSFERASE#)

FILE 'BIOTECHDS'

1250 ARGININE

345 METHYLTRANSFERASE#

L4 1 ARGININE METHYLTRANSFERASE#  
(ARGININE (W) METHYLTRANSFERASE#)

FILE 'BIOSIS'

61402 ARGININE

9218 METHYLTRANSFERASE#

L5 44 ARGININE METHYLTRANSFERASE#  
(ARGININE (W) METHYLTRANSFERASE#)

FILE 'EMBASE'

46360 "ARGININE"

9178 METHYLTRANSFERASE#

L6 66 ARGININE METHYLTRANSFERASE#  
("ARGININE" (W) METHYLTRANSFERASE#)

FILE 'HCAPLUS'

83188 ARGININE

11288 METHYLTRANSFERASE#

L7 79 ARGININE METHYLTRANSFERASE#  
(ARGININE (W) METHYLTRANSFERASE#)

FILE 'NTIS'

288 ARGININE  
35 METHYLTRANSFERASE#  
L8 0 ARGININE METHYLTRANSFERASE#  
(ARGININE (W) METHYLTRANSFERASE#)

FILE 'ESBIOBASE'  
16076 ARGININE  
2647 METHYLTRANSFERASE#  
L9 32 ARGININE METHYLTRANSFERASE#  
(ARGININE (W) METHYLTRANSFERASE#)

FILE 'BIOTECHNO'  
14229 ARGININE  
3989 METHYLTRANSFERASE#  
L10 43 ARGININE METHYLTRANSFERASE#  
(ARGININE (W) METHYLTRANSFERASE#)

FILE 'WPIDS'  
4144 ARGININE  
194 METHYLTRANSFERASE#  
L11 2 ARGININE METHYLTRANSFERASE#  
(ARGININE (W) METHYLTRANSFERASE#)

TOTAL FOR ALL FILES  
L12 400 ARGININE METHYLTRANSFERASE#

=> s l12 and (gene/q or mouse or murine)

FILE 'MEDLINE'  
210763 MOUSE  
89759 MURINE  
L13 35 L1 AND (GENE/Q OR MOUSE OR MURINE)

FILE 'SCISEARCH'  
219978 MOUSE  
92269 MURINE  
L14 37 L2 AND (GENE/Q OR MOUSE OR MURINE)

FILE 'LIFESCI'  
84682 MOUSE  
41177 MURINE  
L15 22 L3 AND (GENE/Q OR MOUSE OR MURINE)

FILE 'BIOTECHDS'  
18663 MOUSE  
1364 MURINE  
L16 1 L4 AND (GENE/Q OR MOUSE OR MURINE)

FILE 'BIOSIS'  
586309 MOUSE  
120340 MURINE  
L17 32 L5 AND (GENE/Q OR MOUSE OR MURINE)

FILE 'EMBASE'  
461645 MOUSE  
83778 MURINE  
L18 43 L6 AND (GENE/Q OR MOUSE OR MURINE)

FILE 'HCAPLUS'  
242778 MOUSE

82230 MURINE  
L19 43 L7 AND (GENE/Q OR MOUSE OR MURINE)

FILE 'NTIS'

3632 MOUSE  
-831 MURINE  
L20 0 L8 AND (GENE/Q OR MOUSE OR MURINE)

FILE 'ESBIOBASE'

63068 MOUSE  
28856 MURINE  
L21 22 L9 AND (GENE/Q OR MOUSE OR MURINE)

FILE 'BIOTECHNO'

181746 MOUSE  
47727 MURINE  
L22 30 L10 AND (GENE/Q OR MOUSE OR MURINE)

FILE 'WPIDS'

13615 MOUSE  
2010 MURINE  
L23 1 L11 AND (GENE/Q OR MOUSE OR MURINE)

TOTAL FOR ALL FILES

L24 266 L12 AND (GENE/Q OR MOUSE OR MURINE)

=> s (steroid or glucocorticoid) (w) receptor#

FILE 'MEDLINE'

64982 STEROID  
19511 GLUCOCORTICOID  
510798 RECEPTOR#  
L25 8873 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

FILE 'SCISEARCH'

46748 STEROID  
19635 GLUCOCORTICOID  
528061 RECEPTOR#  
L26 11582 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

FILE 'LIFESCI'

10489 STEROID  
5093 GLUCOCORTICOID  
173350 RECEPTOR#  
L27 2854 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

FILE 'BIOTECHDS'

2021 STEROID  
212 GLUCOCORTICOID  
8587 RECEPTOR#  
L28 131 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

FILE 'BIOSIS'

75066 STEROID  
22845 GLUCOCORTICOID  
601962 RECEPTOR#  
L29 10689 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

FILE 'EMBASE'

78988 STEROID

29065 GLUCOCORTICOID  
574638 RECEPTOR#  
L30 10565 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

FILE 'HCAPLUS'  
86246 STEROID  
22195 GLUCOCORTICOID  
520156 RECEPTOR#  
L31 10888 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

FILE 'NTIS'  
538 STEROID  
101 GLUCOCORTICOID  
5282 RECEPTOR#  
L32 61 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

FILE 'ESBIOBASE'  
14003 STEROID  
5447 GLUCOCORTICOID  
176734 RECEPTOR#  
L33 3086 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

FILE 'BIOTECHNO'  
15848 STEROID  
7770 GLUCOCORTICOID  
165803 RECEPTOR#  
L34 4460 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

FILE 'WPIDS'  
6368 STEROID  
745 GLUCOCORTICOID  
27056 RECEPTOR#  
L35 202 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

TOTAL FOR ALL FILES  
L36 63391 (STEROID OR GLUCOCORTICOID) (W) RECEPTOR#

=> s transcription?(10a) (activat? or coactivat?)

FILE 'MEDLINE'  
178258 TRANSCRIPTION?  
470972 ACTIVAT?  
2149 COACTIVAT?  
L37 26684 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

FILE 'SCISEARCH'  
151275 TRANSCRIPTION?  
542620 ACTIVAT?  
2966 COACTIVAT?  
L38 30177 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

FILE 'LIFESCI'  
82980 TRANSCRIPTION?  
166374 ACTIVAT?  
1276 COACTIVAT?  
L39 17797 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

FILE 'BIOTECHDS'  
9484 TRANSCRIPTION?  
16193 ACTIVAT?

15 COACTIVAT?  
L40 750 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

FILE 'BIOSIS'

183345 TRANSCRIPTION?  
532844 ACTIVAT?  
2255 COACTIVAT?  
L41 30083 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

FILE 'EMBASE'

151019 TRANSCRIPTION?  
454821 ACTIVAT?  
2072 COACTIVAT?  
L42 24630 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

FILE 'HCAPLUS'

190478 TRANSCRIPTION?  
907705 ACTIVAT?  
2745 COACTIVAT?  
L43 36419 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

FILE 'NTIS'

1959 TRANSCRIPTION?  
26383 ACTIVAT?  
42 COACTIVAT?  
L44 161 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

FILE 'ESBIOBASE'

82361 TRANSCRIPTION?  
175017 ACTIVAT?  
1544 COACTIVAT?  
L45 17461 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

FILE 'BIOTECHNO'

124120 TRANSCRIPTION?  
180986 ACTIVAT?  
1291 COACTIVAT?  
L46 20059 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

FILE 'WPIDS'

7443 TRANSCRIPTION?  
189402 ACTIVAT?  
219 COACTIVAT?  
L47 747 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

TOTAL FOR ALL FILES

L48 204968 TRANSCRIPTION?(10A) (ACTIVAT? OR COACTIVAT?)

=> s (l36 or l48) and methyltransferase#

FILE 'MEDLINE'

12238 METHYLTRANSFERASE#  
L49 129 (L25 OR L37) AND METHYLTRANSFERASE#

FILE 'SCISEARCH'

8352 METHYLTRANSFERASE#  
L50 141 (L26 OR L38) AND METHYLTRANSFERASE#

FILE 'LIFESCI'

3436 METHYLTRANSFERASE#

L51 58 (L27 OR L39) AND METHYLTRANSFERASE#

FILE 'BIOTECHDS'

345 METHYLTRANSFERASE#

L52 0 (L28 OR L40) AND METHYLTRANSFERASE#

FILE 'BIOSIS'

9218 METHYLTRANSFERASE#

L53 109 (L29 OR L41) AND METHYLTRANSFERASE#

FILE 'EMBASE'

9178 METHYLTRANSFERASE#

L54 112 (L30 OR L42) AND METHYLTRANSFERASE#

FILE 'HCAPLUS'

11288 METHYLTRANSFERASE#

L55 146 (L31 OR L43) AND METHYLTRANSFERASE#

FILE 'NTIS'

35 METHYLTRANSFERASE#

L56 0 (L32 OR L44) AND METHYLTRANSFERASE#

FILE 'ESBIOBASE'

2647 METHYLTRANSFERASE#

L57 56 (L33 OR L45) AND METHYLTRANSFERASE#

FILE 'BIOTECHNO'

3989 METHYLTRANSFERASE#

L58 78 (L34 OR L46) AND METHYLTRANSFERASE#

FILE 'WPIDS'

194 METHYLTRANSFERASE#

L59 0 (L35 OR L47) AND METHYLTRANSFERASE#

TOTAL FOR ALL FILES

L60 829 (L36 OR L48) AND METHYLTRANSFERASE#

=> s (l36 or l48) (10a)methyltransferase#

FILE 'MEDLINE'

12238 METHYLTRANSFERASE#

L61 17 (L25 OR L37) (10A)METHYLTRANSFERASE#

FILE 'SCISEARCH'

8352 METHYLTRANSFERASE#

L62 19 (L26 OR L38) (10A)METHYLTRANSFERASE#

FILE 'LIFESCI'

3436 METHYLTRANSFERASE#

L63 14 (L27 OR L39) (10A)METHYLTRANSFERASE#

FILE 'BIOTECHDS'

345 METHYLTRANSFERASE#

L64 0 (L28 OR L40) (10A)METHYLTRANSFERASE#

FILE 'BIOSIS'

9218 METHYLTRANSFERASE#

L65 23 (L29 OR L41) (10A)METHYLTRANSFERASE#

FILE 'EMBASE'

9178 METHYLTRANSFERASE#  
L66 17 (L30 OR L42) (10A) METHYLTRANSFERASE#

FILE 'HCAPLUS'  
11288 METHYLTRANSFERASE#  
L67 33 (L31 OR L43) (10A) METHYLTRANSFERASE#

FILE 'NTIS'  
35 METHYLTRANSFERASE#  
L68 0 (L32 OR L44) (10A) METHYLTRANSFERASE#

FILE 'ESBIOBASE'  
2647 METHYLTRANSFERASE#  
L69 9 (L33 OR L45) (10A) METHYLTRANSFERASE#

FILE 'BIOTECHNO'  
3989 METHYLTRANSFERASE#  
L70 8 (L34 OR L46) (10A) METHYLTRANSFERASE#

FILE 'WPIDS'  
194 METHYLTRANSFERASE#  
L71 0 (L35 OR L47) (10A) METHYLTRANSFERASE#

TOTAL FOR ALL FILES  
L72 140 (L36 OR L48) (10A) METHYLTRANSFERASE#

=> s (124 or 172) not 2000-2001/py

FILE 'MEDLINE'  
548588 2000-2001/PY  
L73 32 (L13 OR L61) NOT 2000-2001/PY

FILE 'SCISEARCH'  
1314665 2000-2001/PY  
L74 35 (L14 OR L62) NOT 2000-2001/PY

FILE 'LIFESCI'  
105420 2000-2001/PY  
L75 24 (L15 OR L63) NOT 2000-2001/PY

FILE 'BIOTECHDS'  
17561 2000-2001/PY  
L76 0 (L16 OR L64) NOT 2000-2001/PY

FILE 'BIOSIS'  
675341 2000-2001/PY  
L77 36 (L17 OR L65) NOT 2000-2001/PY

FILE 'EMBASE'  
590226 2000-2001/PY  
L78 39 (L18 OR L66) NOT 2000-2001/PY

FILE 'HCAPLUS'  
1312271 2000-2001/PY  
L79 45 (L19 OR L67) NOT 2000-2001/PY

FILE 'NTIS'  
0 2000-2001/PY  
L80 0 (L20 OR L68) NOT 2000-2001/PY



FILE 'ESBIOBASE'  
378889 2000-2001/PY  
L81 16 (L21 OR L69) NOT 2000-2001/PY

FILE 'BIOTECHNO'  
160069 2000-2001/PY  
L82 22 (L22 OR L70) NOT 2000-2001/PY

FILE 'WPIDS'  
1212346 2000-2001/PY  
L83 0 (L23 OR L71) NOT 2000-2001/PY

TOTAL FOR ALL FILES  
L84 249 (L24 OR L72) NOT 2000-2001/PY

=> dup rem l84  
PROCESSING COMPLETED FOR L84  
L85 76 DUP REM L84 (173 DUPLICATES REMOVED)

=> d tot

L85 ANSWER 1 OF 76 MEDLINE  
TI How chromatin changes its shape.  
SO SCIENCE, (1999 Aug 20) 285 (5431) 1200-1, 1203.  
Journal code: UJ7; 0404511. ISSN: 0036-8075.  
AU Hagmann M  
AN 1999408060 MEDLINE

L85 ANSWER 2 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
TI Mammalian gene CRX characterization and applications for the detection and treatment of retinal degenerative disease  
SO PCT Int. Appl., 147 pp.  
CODEN: PIXXD2  
IN Freund, Carol L.; McInnes, Roderick R.; Looser, Jens; Cepko, Constance L.; Furukawa, Takahisa; Morrow, Eric M.  
AN 1999:350679 HCAPLUS  
DN 131:1463

| PATENT NO.                                                                 | KIND | DATE     | APPLICATION NO. | DATE     |
|----------------------------------------------------------------------------|------|----------|-----------------|----------|
| WO 9925721                                                                 | A1   | 19990527 | WO 1998-US24322 | 19981113 |
| W: AU, CA, JP, NZ, US                                                      |      |          |                 |          |
| RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE |      |          |                 |          |
| AU 9914089                                                                 | A1   | 19990607 | AU 1999-14089   | 19981113 |

L85 ANSWER 3 OF 76 MEDLINE DUPLICATE 1  
TI Inhibition of DNA **methyltransferase** stimulates the expression of signal transducer and **activator** of **transcription** 1, 2, and 3 genes in colon tumor cells.  
SO PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1999 Nov 23) 96 (24) 14007-12.  
Journal code: PV3; 7505876. ISSN: 0027-8424.  
AU Karpf A R; Peterson P W; Rawlins J T; Dalley B K; Yang Q; Albertsen H; Jones D A  
AN 2000040667 MEDLINE

L85 ANSWER 4 OF 76 MEDLINE DUPLICATE 2  
TI Unusual sites of arginine methylation in Poly(A)-binding protein II and in vitro methylation by protein **arginine methyltransferases**

PRMT1 and PRMT3.

SO JOURNAL OF BIOLOGICAL CHEMISTRY, (1999 May 7) 274 (19) 13229-34.

Journal code: HIV; 2985121R. ISSN: 0021-9258.

AU Smith J J; Rucknagel K P; Schierhorn A; Tang J; Nemeth A; Linder M;  
Herschman H R; Wahle E

AN 1999240708 MEDLINE

L85 ANSWER 5 OF 76 HCAPLUS COPYRIGHT 2001 ACS

TI Analysis of mice carrying targeted mutations of the glucocorticoid  
receptor gene argues against an essential role of glucocorticoid  
signalling for generating adrenal chromaffin cells

SO Development (Cambridge, U. K.) (1999), 126(13), 2935-2944

CODEN: DEVPED; ISSN: 0950-1991

AU Finotto, Susetta; Krieglstein, Kerstin; Schober, Andreas; Deimling,  
Frauke; Lindner, Karin; Bruhl, Barbara; Beier, Konstantin; Metz, Jorgen;  
Garcia-Ararras, Jose E.; Roig-Lopez, Jose L.; Monaghan, Paula; Schmid,  
Wolfgang; Cole, Timothy J.; Kellendonk, Christoph; Tronche, Francois;  
Schutz, Gunther; Unsicker, Klaus

AN 1999:492464 HCAPLUS

DN 131:252700

L85 ANSWER 6 OF 76 MEDLINE

DUPLICATE 3

TI Regulation of transcription by a protein methyltransferase.

SO SCIENCE, (1999 Jun 25) 284 (5423) 2174-7.

Journal code: UJ7; 0404511. ISSN: 0036-8075.

AU Chen D; Ma H; Hong H; Koh S S; Huang S M; Schurter B T; Aswad D W;  
Stallcup M R

AN 1999316081 MEDLINE

L85 ANSWER 7 OF 76 MEDLINE

DUPLICATE 4

TI S-Adenosylmethionine-dependent methylation in *Saccharomyces cerevisiae*.  
Identification of a novel protein arginine  
methyltransferase.

SO JOURNAL OF BIOLOGICAL CHEMISTRY, (1999 Jan 8) 274 (2) 814-24.

Journal code: HIV; 2985121R. ISSN: 0021-9258.

AU Niewmierzycka A; Clarke S

AN 1999091619 MEDLINE

L85 ANSWER 8 OF 76 HCAPLUS COPYRIGHT 2001 ACS

TI Activation of human O6-methylguanine-DNA methyltransferase gene by  
glucocorticoid hormone

SO Oncogene (1999), 18(2), 525-532

CODEN: ONCNE5; ISSN: 0950-9232

AU Biswas, Tapan; Ramana, Chilakamarti V.; Srinivasan, Ganesan; Boldogh,  
Istvan; Hazra, Tapas K.; Chen, Zhenping; Tano, Keizo; Thompson, E. Brad;  
Mitra, Sankar

AN 1999:117843 HCAPLUS

DN 130:276964

L85 ANSWER 9 OF 76 SCISEARCH COPYRIGHT 2001 ISI (R)

TI RNase treatment of yeast and mammalian cell extracts affects in vitro  
substrate methylation by type I protein arginine N-methyltransferases

SO BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, (7 JUN 1999) Vol.  
259, No. 2, pp. 391-400.

Publisher: ACADEMIC PRESS INC, 525 B ST, STE 1900, SAN DIEGO, CA  
92101-4495.

ISSN: 0006-291X.

AU Frankel A; Clarke S (Reprint)

AN 1999:478606 SCISEARCH

- L85 ANSWER 10 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.  
 TI [Mechanisms of BTG2 activity, a transcriptional target of p53: Evidences and hypothesis].  
 MECANISMES D'ACTION DE BTG2, **GENE** CIBLE DE P53: DONNEES ACQUISES ET HYPOTHESES.  
 SO Bulletin du Cancer, (1999) 86/4 (358-364).  
 Refs: 42  
 ISSN: 0007-4551 CODEN: BUCABS  
 AU Puisieux A.; Magaud J.-P.  
 AN 1999174896 EMBASE
- L85 ANSWER 11 OF 76 MEDLINE DUPLICATE 5  
 TI Arginine methylation and binding of Hrp1p to the efficiency element for mRNA 3'-end formation.  
 SO RNA, (1999 Feb) 5 (2) 272-80.  
 Journal code: CHB; 9509184. ISSN: 1355-8382.  
 AU Valentini S R; Weiss V H; Silver P A  
 AN 1999146783 MEDLINE
- L85 ANSWER 12 OF 76 SCISEARCH COPYRIGHT 2001 ISI (R)  
 TI Identification and characterization of a novel mammalian CpG binding **transcriptional activator** that shares a motif with DNA **methyltransferase** and HRX proteins.  
 SO EXPERIMENTAL HEMATOLOGY, (JUL 1999) Vol: 27, No. 7, Supp: [1], pp. 61-61.  
 Publisher: ELSEVIER SCIENCE INC, 655 AVENUE OF THE AMERICAS, NEW YORK, NY 10010.  
 ISSN: 0301-472X.  
 AU Voo K S (Reprint); Carlone D L; Jacobsen B M; Flodin A; Skalnik D G  
 AN 1999:528880 SCISEARCH
- L85 ANSWER 13 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI Neural regulation of phenylethanolamine N-methyltransferase (PNMT) gene expression in bovine chromaffin cells differs from other catecholamine enzyme genes.  
 SO J. Mol. Neurosci. (1999), 12(1), 53-68  
 CODEN: JMNEES; ISSN: 0895-8696  
 AU Lee, Ying-Shuan-Eda; Raia, Gabrielle; Tonshoff, Christianne; Evinger, Marian J.  
 AN 1999:592097 HCAPLUS  
 DN 131:347371
- L85 ANSWER 14 OF 76 MEDLINE DUPLICATE 6  
 TI delta-N-methylarginine is a novel posttranslational modification of arginine residues in yeast proteins.  
 SO JOURNAL OF BIOLOGICAL CHEMISTRY, (1998 Nov 6) 273 (45) 29283-6.  
 Journal code: HIV; 2985121R. ISSN: 0021-9258.  
 AU Zobel-Thropp P; Gary J D; Clarke S  
 AN 1999009026 MEDLINE
- L85 ANSWER 15 OF 76 MEDLINE DUPLICATE 7  
 TI Identification of protein-arginine N-methyltransferase as 10-formyltetrahydrofolate dehydrogenase.  
 SO JOURNAL OF BIOLOGICAL CHEMISTRY, (1998 Oct 16) 273 (42) 27374-82.  
 Journal code: HIV; 2985121R. ISSN: 0021-9258.  
 AU Kim S; Park G H; Joo W A; Paik W K; Cook R J; Williams K R  
 AN 1998438510 MEDLINE
- L85 ANSWER 16 OF 76 MEDLINE DUPLICATE 8

TI PRMT 3, a type I protein arginine N-methyltransferase that differs from  
 PRMT1 in its oligomerization, subcellular localization, substrate  
 specificity, and regulation.  
 SO JOURNAL OF BIOLOGICAL CHEMISTRY, (1998 Jul 3) 273 (27) 16935-45.  
 Journal code: HIV; 2985121R. ISSN: 0021-9258.  
 AU Tang J; Gary J D; Clarke S; Herschman H R  
 AN 1998307932 MEDLINE

L85 ANSWER 17 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI C-Fos deficiency inhibits induction of mRNA for some, but not all,  
 neurotransmitter biosynthetic enzymes by immobilization stress  
 SO J. Neurochem. (1998), 70(5), 1935-1940  
 CODEN: JONRA9; ISSN: 0022-3042  
 AU Serova, Lidia I.; Saez, Enrique; Spiegelman, Bruce M.; Sabban, Esther L.  
 AN 1998:270959 HCAPLUS  
 DN 129:26592

L85 ANSWER 18 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.DUPLICATE 9  
 TI Enzymatic methylation of recombinant TIS21 protein-arginine residues.  
 SO Biochemistry and Molecular Biology International, (1998) 45/5 (871-878).  
 Refs: 26  
 ISSN: 1039-9712 CODEN: BMBIES  
 AU Lim I.K.; Park T.-J.; Kim S.; Lee H.W.; Paik W.K.  
 AN 1998311260 EMBASE

L85 ANSWER 19 OF 76 MEDLINE DUPLICATE 10  
 TI p53 is involved in regulation of the DNA repair gene O6-methylguanine-DNA  
 methyltransferase (MGMT) by DNA damaging agents.  
 SO ONCOGENE, (1998 Aug 20) 17 (7) 845-51.  
 Journal code: ONC; 8711562. ISSN: 0950-9232.  
 AU Grombacher T; Eichhorn U; Kaina B  
 AN 1998451281 MEDLINE

L85 ANSWER 20 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI The 4S benzo(a)pyrene-binding protein is not a transcriptional activator  
 of Cyp1a1 gene in Ah receptor-deficient (AHR -/-) transgenic mice  
 SO Arch. Biochem. Biophys. (1998), 349(2), 349-355  
 CODEN: ABBIA4; ISSN: 0003-9861  
 AU Foussat, Julie; Costet, Philippe; Galtier, Pierre; Pineau, Thierry; Lesca,  
 Pierre  
 AN 1998:53563 HCAPLUS  
 DN 128:163820

L85 ANSWER 21 OF 76 MEDLINE DUPLICATE 11  
 TI Identification and characterization of two putative human **arginine  
 methyltransferases** (HRMT1L1 and HRMT1L2).  
 SO GENOMICS, (1998 Mar 15) 48 (3) 330-40.  
 Journal code: GEN; 8800135. ISSN: 0888-7543.  
 AU Scott H S; Antonarakis S E; Lalioti M D; Rossier C; Silver P A; Henry M F  
 AN 1998207248 MEDLINE

L85 ANSWER 22 OF 76 MEDLINE DUPLICATE 12  
 TI Phenylethanolamine N-methyltransferase gene expression: synergistic  
 activation by Egr-1, AP-2 and the glucocorticoid receptor.  
 SO BRAIN RESEARCH. MOLECULAR BRAIN RESEARCH, (1998 Oct 30) 61 (1-2) 154-61.  
 Journal code: MBR; 8908640. ISSN: 0169-328X.  
 AU Wong D L; Siddall B J; Ebert S N; Bell R A; Her S  
 AN 1999013984 MEDLINE

L85 ANSWER 23 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI Neural control of phenylethanolamine-N-methyltransferase via cholinergic activation of Egr-I  
 SO Adv. Pharmacol. (San Diego) (1998), 42(Catecholamines), 77-81  
 CODEN: ADPHEL; ISSN: 1054-3589  
 AU Wong, Dona L.; Ebert, Steven N.; Morita, Kyoji  
 AN 1997:711960 HCAPLUS  
 DN 128:631

L85 ANSWER 24 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI RNA and protein interactions modulated by protein arginine methylation  
 SO Prog. Nucleic Acid Res. Mol. Biol. (1998), 61, 65-131  
 CODEN: PNMBAF; ISSN: 0079-6603  
 AU Gary, Jonathan D.; Clarke, Steven  
 AN 1999:1415 HCAPLUS  
 DN 130:193393

L85 ANSWER 25 OF 76 MEDLINE DUPLICATE 13  
 TI Protein N-arginine methylation in adenosine dialdehyde-treated lymphoblastoid cells.  
 SO ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS, (1998 Mar 1) 351 (1) 53-9.  
 Journal code: 6SK; 0372430. ISSN: 0003-9861.  
 AU Li C; Ai L S; Lin C H; Hsieh M; Li Y C; Li S Y  
 AN 1998162641 MEDLINE

L85 ANSWER 26 OF 76 MEDLINE DUPLICATE 14  
 TI Construction of a 2.5-Mb integrated physical and **gene** map of distal 21q22.3.  
 SO GENOMICS, (1998 Apr 1) 49 (1) 1-13.  
 Journal code: GEN; 8800135. ISSN: 0888-7543.  
 AU Lapenta V; Sossi V; Gosset P; Vayssettes C; Vitali T; Rabatel N; Tassone F; Blouin J L; Scott H S; Antonarakis S E; Creau N; Brahe C  
 AN 1998234538 MEDLINE

L85 ANSWER 27 OF 76 MEDLINE DUPLICATE 15  
 TI Identification of N(G)-methylarginine residues in human heterogeneous RNP protein A1: Phe/Gly-Gly-Gly-Arg-Gly-Gly-Gly/Phe is a preferred recognition motif.  
 SO BIOCHEMISTRY, (1997 Apr 29) 36 (17) 5185-92.  
 Journal code: A0G; 0370623. ISSN: 0006-2960.  
 AU Kim S; Merrill B M; Rajpurohit R; Kumar A; Stone K L; Papov V V; Schneiders J M; Szer W; Wilson S H; Paik W K; Williams K R  
 AN 97282571 MEDLINE

L85 ANSWER 28 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI A putative leucine zipper activator of Pasteurella haemolytica leukotoxin transcription and the potential for modulation of its synthesis by slipped-strand mispairing  
 SO Infect. Immun. (1997), 65(9), 3970-3975  
 CODEN: INFIBR; ISSN: 0019-9567  
 AU Highlander, Sarah K.; Hang, Vinh T.  
 AN 1997:596469 HCAPLUS  
 DN 127:315461

L85 ANSWER 29 OF 76 MEDLINE DUPLICATE 16  
 TI Detection of weakly expressed genes in the rostral ventrolateral medulla of the rat using micropunch and reverse transcription-polymerase chain reaction techniques.  
 SO CLINICAL AND EXPERIMENTAL PHARMACOLOGY AND PHYSIOLOGY, (1997 Sep-Oct) 24

(9-10) 755-9.

Journal code: DD8; 0425076. ISSN: 0305-1870.

AU Comer A M; Yip S; Lipski J  
AN 97461017 MEDLINE

L85 ANSWER 30 OF 76 SCISEARCH COPYRIGHT 2001 ISI (R)  
TI Biological methylation of myelin basic protein: Enzymology and biological significance  
SO INTERNATIONAL JOURNAL OF BIOCHEMISTRY & CELL BIOLOGY, (MAY 1997) Vol. 29, No. 5, pp. 743-751.  
Publisher: PERGAMON-ELSEVIER SCIENCE LTD, THE BOULEVARD, LANGFORD LANE, KIDLINGTON, OXFORD, ENGLAND OX5 1GB.  
ISSN: 1357-2725.

AU Kim S; Lim I K; Park G H; Paik W K (Reprint)  
AN 97:582049 SCISEARCH

L85 ANSWER 31 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
TI Identification and mapping of a novel human **gene**, HRMT1L1, homologous to the rat protein arginine N-methyltransferase 1 (PRMT1) **gene**

SO Mamm. Genome (1997), 8(7), 526-529  
CODEN: MAMGEC; ISSN: 0938-8990

AU Katsanis, Nicholas; Yaspo, Marie-Laure; Fisher, Elizabeth M.C.  
AN 1997:466667 HCAPLUS  
DN 127:131780

L85 ANSWER 32 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.DUPLICATE 17  
TI Human erythrocyte protein L-isoaspartyl methyltransferase: Heritability of basal activity and genetic polymorphism for thermal stability.  
SO Archives of Biochemistry and Biophysics, (1997) 346/2 (277-286).

Refs: 47

ISSN: 0003-9861 CODEN: ABBIA4

AU David C.L.; Szumlanski C.L.; DeVry C.G.; Park-Hah J.O.; Clarke S.; Weinshilboum R.M.; Aswad D.W.  
AN 97315267 EMBASE

L85 ANSWER 33 OF 76 MEDLINE DUPLICATE 18  
TI A protein-**arginine methyltransferase** binds to the intracytoplasmic domain of the IFNAR1 chain in the type I interferon receptor.

SO EMBO JOURNAL, (1997 Jan 15) 16 (2) 260-6.  
Journal code: EMB; 8208664. ISSN: 0261-4189.

AU Abramovich C; Yakobson B; Chebath J; Revel M  
AN 97180929 MEDLINE

L85 ANSWER 34 OF 76 MEDLINE DUPLICATE 19  
TI The mammalian immediate-early TIS21 protein and the leukemia-associated BTG1 protein interact with a protein-arginine N-methyltransferase.

SO JOURNAL OF BIOLOGICAL CHEMISTRY, (1996 Jun 21) 271 (25) 15034-44.  
Journal code: HIV; 2985121R. ISSN: 0021-9258.

AU Lin W J; Gary J D; Yang M C; Clarke S; Herschman H R  
AN 96278999 MEDLINE

L85 ANSWER 35 OF 76 MEDLINE DUPLICATE 20  
TI The essential yeast RNA binding protein Np13p is methylated.  
SO PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1996 Nov 26) 93 (24) 13641-6.

Journal code: PV3; 7505876. ISSN: 0027-8424.

AU Siebel C W; Guthrie C

AN 97098447 MEDLINE

L85 ANSWER 36 OF 76 MEDLINE DUPLICATE 21

TI The predominant protein-**arginine methyltransferase** from *Saccharomyces cerevisiae*.

SO JOURNAL OF BIOLOGICAL CHEMISTRY, (1996 May 24) 271 (21) 12585-94.  
Journal code: HIV; 2985121R. ISSN: 0021-9258.

AU Gary J D; Lin W J; Yang M C; Herschman H R; Clarke S

AN 96218186 MEDLINE

L85 ANSWER 37 OF 76 MEDLINE DUPLICATE 22

TI Human breast cancer progression can be regulated by dominant trans-acting factors in somatic cell hybridization studies.

SO CANCER RESEARCH, (1996 Aug 1) 56 (15) 3560-9.  
Journal code: CNF; 2984705R. ISSN: 0008-5472.

AU Safarians S; Sternlicht M D; Yamanishi D T; Love S M; Barsky S H

AN 96357182 MEDLINE

---

L85 ANSWER 38 OF 76 HCAPLUS COPYRIGHT 2001 ACS

TI Glucocorticoid control of phenylethanolamine N-methyltransferase gene expression: Implications for stress and disorders of the stress axis

SO Stress, Proc. Int. Symp. Catecholamines Other Neurotransm. Stress, 6th (1996), Meeting Date 1995, Volume 2, 677-693. Editor(s): McCarty, Richard. Publisher: Harwood, Amsterdam, Neth.  
CODEN: 63WCA9

AU Wong, D. L.; Ebert, S. N.; Morita, K.

AN 1997:29975 HCAPLUS

DN 126:99505

L85 ANSWER 39 OF 76 MEDLINE DUPLICATE 23

TI Stress responses in alfalfa (*Medicago sativa* L.). XX. Transcriptional activation of phenylpropanoid pathway genes in elicitor-induced cell suspension cultures.

SO PLANT MOLECULAR BIOLOGY, (1996 Feb) 30 (3) 427-38.  
Journal code: A60; 9106343. ISSN: 0167-4412.

AU Ni W; Fahrendorf T; Ballance G M; Lamb C J; Dixon R A

AN 96189259 MEDLINE

L85 ANSWER 40 OF 76 HCAPLUS COPYRIGHT 2001 ACS

TI Regulation of the *Serratia marcescens* extracellular nuclease: positive control by a homolog of P2 Ogr encoded by a cryptic prophage

SO J. Mol. Biol. (1996), 256(2), 264-78  
CODEN: JMOBAK; ISSN: 0022-2836

AU Jin, Shida; Chen, Yi-chi; Christie, Gail E.; Benedik, Michael J.

AN 1996:120396 HCAPLUS

DN 124:195763

L85 ANSWER 41 OF 76 BIOSIS COPYRIGHT 2001 BIOSIS

TI Isolation and mapping of two human **genes** homologues (hHMT1 and hHMT2) to a yeast **arginine methyltransferase**.

SO American Journal of Human Genetics, (1996) Vol. 59, No. 4 SUPPL., pp. A159.  
Meeting Info.: 46th Annual Meeting of the American Society of Human Genetics San Francisco, California, USA October 29-November 2, 1996  
ISSN: 0002-9297.

AU Scott, H. S. (1); Henry, M. F.; Lalioti, M. D. (1); Rossier, C. (1); Silver, P. A.; Antonarakis, S. E. (1)

AN 1996:555880 BIOSIS

L85 ANSWER 42 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI Differential activation of the rat phenylethanolamine N-methyltransferase gene by Sp1 and Egr-1  
 SO J. Biol. Chem. (1995), 270(29), 17299-305  
 CODEN: JBCHA3; ISSN: 0021-9258  
 AU Ebert, Steven N.; Wong, Dona L.  
 AN 1995:707848 HCAPLUS  
 DN 123:162354

L85 ANSWER 43 OF 76 MEDLINE DUPLICATE 24  
 TI Transcriptional activation of the Epstein-Barr virus latency C promoter after 5-azacytidine treatment: evidence that demethylation at a single CpG site is crucial.  
 SO MOLECULAR AND CELLULAR BIOLOGY, (1995 Nov) 15 (11) 6150-9.  
 Journal code: NGY; 8109087. ISSN: 0270-7306.  
 AU Robertson K D; Hayward S D; Ling P D; Samid D; Ambinder R F  
 AN 96025995 MEDLINE

L85 ANSWER 44 OF 76 SCISEARCH COPYRIGHT 2001 ISI (R) DUPLICATE 25  
 TI **TRANSCRIPTIONAL ACTIVATION OF THE PHENYLETHANOLAMINE N-METHYLTRANSFERASE GENE BY EGR-1 AND SP1**  
 SO JOURNAL OF NEUROCHEMISTRY, (1995) Vol. 65, Supp. S, pp. S77.  
 ISSN: 0022-3042.  
 AU WONG D L (Reprint); EBERT S N  
 AN 95:640264 SCISEARCH

L85 ANSWER 45 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V. DUPLICATE 26  
 TI Structural specificity of substrate for S-adenosylmethionine:protein arginine N-methyltransferases.  
 SO Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology, (1995) 1248/1 (11-18).  
 ISSN: 0167-4838 CODEN: BBAEDZ  
 AU Rawal N.; Rajpurohit R.; Lischwe M.A.; Williams K.R.; Paik W.K.; Kim S.  
 AN 95103187 EMBASE

L85 ANSWER 46 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI Egr-1 activation of rat adrenal phenylethanolamine N-methyltransferase gene  
 SO J. Biol. Chem. (1994), 269(33), 20885-98  
 CODEN: JBCHA3; ISSN: 0021-9258  
 AU Ebert, Steven N.; Balt, Steven L.; Hunter, Joan P. B.; Gashler, Andrea; Sukhatme, Vikas; Wong, Dona L.  
 AN 1994:550482 HCAPLUS  
 DN 121:150482

L85 ANSWER 47 OF 76 MEDLINE DUPLICATE 27  
 TI Peptides with **sequences** similar to glycine, arginine-rich motifs in proteins interacting with RNA are efficiently recognized by methyltransferase(s) modifying arginine in numerous proteins.  
 SO JOURNAL OF BIOLOGICAL CHEMISTRY, (1993 May 15) 268 (14) 10501-9.  
 Journal code: HIV; 2985121R. ISSN: 0021-9258.  
 AU Najbauer J; Johnson B A; Young A L; Aswad D W  
 AN 93252938 MEDLINE

L85 ANSWER 48 OF 76 BIOSIS COPYRIGHT 2001 BIOSIS  
 TI Phenylethanol amine N-methyltransferase gene activation: EGR-1 and **glucocorticoid receptor** synergy.  
 SO Society for Neuroscience Abstracts, (1993) Vol. 19, No. 1-3, pp. 1692.  
 Meeting Info.: 23rd Annual Meeting of the Society for Neuroscience



Washington, D.C., USA November 7-12, 1993  
ISSN: 0190-5295.

AU Wong, D. L.; Ebert, S. N.; Balt, S. L.; Hunter, J. P. B.  
AN 1994:48315 BIOSIS

L85 ANSWER 49 OF 76 SCISEARCH COPYRIGHT 2001 ISI (R)  
TI EGR-1, A TRANSCRIPTIONAL REGULATOR OF PHENYLETHANOLAMINE N-  
**METHYLTRANSFERASE** GENE-EXPRESSION - SYNERGISTIC EFFECTS WITH THE  
**GLUCOCORTICOID RECEPTOR**  
SO NEUROPSYCHOPHARMACOLOGY, (SEP 1993) Vol. 9, No. 2, Supp. S, pp. S83-S84.  
ISSN: 0893-133X.  
AU WONG D L (Reprint); EBERT S N; HUNTER J; GASHER A; SUKHATME V P  
AN 93:542475 SCISEARCH

L85 ANSWER 50 OF 76 MEDLINE DUPLICATE 28  
TI Adrenal phenylethanolamine N-methyltransferase induction in  
relation to **glucocorticoid receptor** dynamics: evidence  
that acute exposure to high cortisol levels is sufficient to induce the  
enzyme.  
SO JOURNAL OF NEUROCHEMISTRY, (1992 May) 58 (5) 1853-62.  
Journal code: JAV; 2985190R. ISSN: 0022-3042.  
AU Betito K; Diorio J; Meaney M J; Boksa P  
AN 92218934 MEDLINE

L85 ANSWER 51 OF 76 SCISEARCH COPYRIGHT 2001 ISI (R)  
TI STUDIES ON N(G)-METHYLARGININE DERIVATIVES IN MYELIN BASIC-PROTEIN FROM  
DEVELOPING AND MUTANT **MOUSE**-BRAIN  
SO BIOCHEMICAL JOURNAL, (01 NOV 1992) Vol. 287, Part 3, pp. 929-935.  
ISSN: 0264-6021.  
AU RAWAL N; LEE Y J; PAIK W K; KIM S (Reprint)  
AN 92:653725 SCISEARCH

L85 ANSWER 52 OF 76 SCISEARCH COPYRIGHT 2001 ISI (R) DUPLICATE 29  
TI A REGION OF THE ADA DNA-REPAIR PROTEIN REQUIRED FOR THE ACTIVATION OF ADA  
TRANSCRIPTION IS NOT NECESSARY FOR ACTIVATION OF ALKA  
SO PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF  
AMERICA, (1991) Vol. 88, No. 20, pp. 9001-9005.  
AU SHEVELL D E; WALKER G C (Reprint)  
AN 91:573137 SCISEARCH

L85 ANSWER 53 OF 76 BIOSIS COPYRIGHT 2001 BIOSIS  
TI MOLECULAR ANALYSIS OF BACILLUS-SUBTILIS ADA MUTANTS DEFICIENT IN THE  
ADAPTIVE RESPONSE TO SIMPLE ALKYLATING AGENTS.  
SO J BACTERIOL, (1991) 173 (24), 7834-7840.  
CODEN: JOBAAY. ISSN: 0021-9193.  
AU MOROHOSHI F; HAYASHI K; MUNAKATA N  
AN 1992:95704 BIOSIS

L85 ANSWER 54 OF 76 SCISEARCH COPYRIGHT 2001 ISI (R)  
TI ENZYMATIC METHYLATION OF MYELIN BASIC-PROTEIN IN MYELIN  
SO BIOCHEMICAL JOURNAL, (1991) Vol. 275, No. APR, pp. 381-387.  
AU GHOSH S K; RAWAL N; SYED S K; PAIK W K; KIM S (Reprint)  
AN 91:244762 SCISEARCH

L85 ANSWER 55 OF 76 SCISEARCH COPYRIGHT 2001 ISI (R) DUPLICATE 30  
TI AN ENZYME-LINKED-IMMUNOSORBENT-ASSAY FOR MYELIN BASIC PROTEIN-SPECIFIC  
PROTEIN METHYLASE-I  
SO JOURNAL OF NEUROSCIENCE METHODS, (1991) Vol. 37, No. 2, pp. 133-140.  
AU RAWAL N; PAIK W K; KIM S (Reprint)

AN 91:296925 SCISEARCH

L85 ANSWER 56 OF 76 BIOSIS COPYRIGHT 2001 BIOSIS  
TI NEURAL AND HORMONAL REGULATION OF THE TYROSINE HYDROXYLASE GENE IN ADRENAL  
MEDULLARY CELLS PARTICIPATION OF C-FOS AND AP1 FACTORS.  
SO MOL CELL NEUROSCI, (1990) 1 (3), 202-213.  
CODEN: MOCNED. ISSN: 1044-7431.  
AU STACHOWIAK M K; GOC A; HONG J S; KAPLAN B B; STACHOWIAK E K  
AN 1992:30788 BIOSIS

L85 ANSWER 57 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.DUPLICATE 31  
TI Substrate specificity for myelin basic protein-specific protein methylase  
I.  
SO Biochimica et Biophysica Acta - Protein Structure and Molecular  
Enzymology, (1990) 1039/2 (142-148).  
ISSN: 0167-4838 CODEN: BBAEDZ  
AU Ghosh S.K.; Syed S.K.; Jung S.; Paik W.K.; Kim S.  
AN 90183956 EMBASE

L85 ANSWER 58 OF 76 BIOSIS COPYRIGHT 2001 BIOSIS  
TI SITE-DIRECTED MUTATION OF THE ESCHERICHIA-COLI ADA GENE EFFECTS OF  
SUBSTITUTION OF METHYL ACCEPTOR CYSTEINE-321 BY HISTIDINE IN ADA PROTEIN.  
SO J BACTERIOL, (1989) 171 (3), 1535-1543.  
CODEN: JOBAAY. ISSN: 0021-9193.  
AU TANO K; BHATTACHARYYA D; FOOTE R S; MURAL R J; MITRA S  
AN 1989:227441 BIOSIS

L85 ANSWER 59 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
TI The glucocorticoid receptor in the adrenal gland is localized in the  
cytoplasm of adrenaline cells  
SO Acta Physiol. Scand. (1989), 137(4), 559-60  
CODEN: APSCAX; ISSN: 0001-6772  
AU Ceccatelli, Sandra; Dagerlind, A.; Schalling, M.; Wikstroem, A. C.; Okret,  
S.; Gustafsson, J. A.; Goldstein, M.; Hoekfelt, T.  
AN 1990:70205 HCAPLUS  
DN 112:70205

L85 ANSWER 60 OF 76 MEDLINE DUPLICATE 32  
TI Glucocorticoid receptor-immunoreactivity in C1, C2, and C3 adrenergic  
neurons that project to the hypothalamus or to the spinal cord in the rat.  
SO JOURNAL OF COMPARATIVE NEUROLOGY, (1989 Jul 1) 285 (1) 107-16.  
Journal code: HUV; 0406041. ISSN: 0021-9967.  
AU Sawchenko P E; Bohn M C  
AN 89327587 MEDLINE

L85 ANSWER 61 OF 76 MEDLINE DUPLICATE 33  
TI Morphometrical and microdensitometrical studies on phenylethanolamine-N-  
**methyltransferase**- and neuropeptide Y-immunoreactive nerve  
terminals and on **glucocorticoid receptor**  
-immunoreactive nerve cell nuclei in the paraventricular hypothalamic  
nucleus in adult and old male rats.  
SO NEUROSCIENCE, (1988 Aug) 26 (2) 479-92.  
Journal code: NZR; 7605074. ISSN: 0306-4522.  
AU Zoli M; Agnati L F; Fuxe K; Zini I; Merlo Pich E; Grimaldi R; Harfstrand  
A; Goldstein M; Wikstrom A C; Gustafsson J A  
AN 89015067 MEDLINE

L85 ANSWER 62 OF 76 BIOSIS COPYRIGHT 2001 BIOSIS  
TI FUNCTIONAL SITES OF THE ADA REGULATORY PROTEIN OF ESCHERICHIA-COLI

ANALYSIS BY AMINO ACID SUBSTITUTIONS.  
SO J MOL BIOL, (1988) 201 (2), 261-272.  
CODEN: JMOBAK. ISSN: 0022-2836.  
AU TAKANO K; NAKABEPPU Y; SEKIGUCHI M  
AN 1988:330879 BIOSIS

L85 ANSWER 63 OF 76 MEDLINE DUPLICATE 34  
TI Glucocorticoid receptor immunoreactivity in monoaminergic neurons of rat brain.  
SO PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1986 Dec) 83 (24) 9779-83.  
Journal code: PV3; 7505876. ISSN: 0027-8424.  
AU Harfstrand A; Fuxe K; Cintra A; Agnati L F; Zini I; Wikstrom A C; Okret S; Yu Z Y; Goldstein M; Steinbusch H; +  
AN 87092331 MEDLINE

L85 ANSWER 64 OF 76 LIFESCI COPYRIGHT 2001 CSA  
TI Glucocorticoid receptor immunoreactivity in monoaminergic neurons of rat brain.  
SO PROC. NATL. ACAD. SCI. USA., (1986) vol. 83, no. 24, pp. 9779-9783.  
AU Haerfstrand, A.; Fuxe, K.; Cintra, A.; Agnati, L.F.; Zini, I.; Wikstroem, A.-C.; Okret, S.; Yu, Z.-Y.; Goldstein, M.; et al..  
AN 86:67751 LIFESCI

L85 ANSWER 65 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.  
TI .beta.-Adrenergic stimulation of protein (**arginine**) **methyltransferase** activity in cultured cerebral cells from embryonic mice.  
SO Journal of Neuroscience Research, (1986) 16/2 (377-386).  
CODEN: JNREDK  
AU Amur S.G.; Shanker G.; Pieringer R.A.  
AN 86233852 EMBASE

L85 ANSWER 66 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.  
TI Correlation between inhibition of myelin basic protein (**arginine**) **methyltransferase** by sinefungin and lack of compact myelin formation in cultures of cerebral cells from embryonic mice.  
SO Journal of Neuroscience Research, (1986) 16/2 (367-376).  
CODEN: JNREDK  
AU Amur S.G.; Shanker G.; Cochran J.M.; et al.  
AN 86233851 EMBASE

L85 ANSWER 67 OF 76 LIFESCI COPYRIGHT 2001 CSA DUPLICATE 35  
TI Myelin basic protein-specific protein methylase I activity in shiverer mutant **mouse** brain.  
SO J. NEUROSCI. RES., (1986) vol. 16, no. 2, pp. 357-365.  
AU Kim, S.; Tuck, M.; Ho, L.-L.; Campagnoni, A.T.; Barbarese, E.; Knobler, R.L.; Lublin, F.D.; Chanderkar, L.P.; Paik, W.K.  
AN 86:70907 LIFESCI

L85 ANSWER 68 OF 76 MEDLINE DUPLICATE 36  
TI **Glucocorticoid receptors** and regulation of phenylethanolamine-N-**methyltransferase** activity in cultured chromaffin cells.  
SO JOURNAL OF NEUROSCIENCE, (1985 Aug) 5 (8) 2161-8.  
Journal code: JDF; 8102140. ISSN: 0270-6474.  
AU Kelner K L; Pollard H B  
AN 85263810 MEDLINE

L85 ANSWER 69 OF 76 MEDLINE DUPLICATE 37  
 TI Morphometrical analysis of the distribution of corticotrophin releasing factor, **glucocorticoid receptor** and phenylethanolamine-N-**methyltransferase** immunoreactive structures in the paraventricular hypothalamic nucleus of the rat.  
 SO NEUROSCIENCE LETTERS, (1985 Mar 15) 54 (2-3) 147-52.  
 Journal code: N7N; 7600130. ISSN: 0304-3940.  
 AU Agnati L F; Fuxe K; Yu Z Y; Harfstrand A; Okret S; Wikstrom A C; Goldstein M; Zoli M; Vale W; Gustafsson J A  
 AN 85188868 MEDLINE

L85 ANSWER 70 OF 76 MEDLINE DUPLICATE 38  
 TI Regulation of myelin basic protein (**arginine**) **methyltransferase** by thyroid hormone in myelinogenic cultures of cells dissociated from embryonic **mouse** brain.  
 SO JOURNAL OF NEUROCHEMISTRY, (1984 Aug) 43 (2) 494-8.  
 Journal code: JAV; 2985190R. ISSN: 0022-3042.  
 AU Amur S G; Shanker G; Pieringer R A  
 AN 84241798 MEDLINE

L85 ANSWER 71 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI Glucocorticoid regulation of phenylethanolamine N-methyltransferase (PNMT) in organ culture of superior cervical ganglia  
 SO Dev. Biol. (1984), 105(1), 130-6  
 CODEN: DEBIAO; ISSN: 0012-1606  
 AU Bohn, M. C.; Bloom, E.; Goldstein, M.; Black, I. B.  
 AN 1984:523374 HCAPLUS  
 DN 101:123374

L85 ANSWER 72 OF 76 MEDLINE DUPLICATE 39  
 TI The relationship of myelin basic protein (**arginine**) **methyltransferase** to myelination in **mouse** spinal cord.  
 SO JOURNAL OF NEUROCHEMISTRY, (1982 Jul) 39 (1) 244-7.  
 Journal code: JAV; 2985190R. ISSN: 0022-3042.  
 AU Crang A J; Jacobson W  
 AN 82215833 MEDLINE

L85 ANSWER 73 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI The methylation in vitro of myelin basic protein by arginine methylase from **mouse** spinal cord  
 SO Biochem. Soc. Trans. (1980), 8(5), 611-12  
 CODEN: BCSTB5; ISSN: 0300-5127  
 AU Crang, A. John; Jacobson, W.  
 AN 1981:78733 HCAPLUS  
 DN 94:78733

L85 ANSWER 74 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.  
 TI Protein methylation in animal cells. I. Purification and properties of S-adenosyl-L-methionine: Protein(**arginine**) n.methyltransferase from krebs II ascites cells.  
 SO Biochimica et Biophysica Acta, (1978) 519/1 (243-254).  
 CODEN: BBACAQ  
 AU Casellas P.; Jeanteur P.J.  
 AN 78375803 EMBASE

L85 ANSWER 75 OF 76 HCAPLUS COPYRIGHT 2001 ACS  
 TI Effects of carcinogens and other agents on histone methylation in rat liver nuclei by endogenous histone lysine methyltransferase  
 SO Cancer Res. (1974), 34(6), 1424-8

CODEN: CNREA8

AU Baxter, C. Stuart; Byvoet, Paul  
AN 1974:459070 HCAPLUS  
DN 81:59070

L85 ANSWER 76 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.  
TI Effects of carcinogens and other agents on histone methylation by a  
histone **arginine methyltransferase** purified from rat  
liver cytoplasm.  
SO Cancer Research, (1974) 34/6 (1418-1423).  
CODEN: CNREA8  
AU Baxter C.S.; Byvoet P.  
AN 75022818 EMBASE

=> d ab 7,11,12,18,24,33,34,45,47,76

L85 ANSWER 7 OF 76 MEDLINE DUPLICATE 4  
AB We used **sequence** motifs conserved in S-adenosylmethionine-  
dependent methyltransferases to identify 26 putative methyltransferases  
from the complete genome of the yeast *Saccharomyces cerevisiae*. Seven  
**sequences** with the best matches to the methyltransferase consensus  
motifs were selected for further study. We prepared yeast disruption  
mutants of each of the **genes** encoding these **sequences**,  
and we found that disruption of the YJL125c **gene** is lethal,  
whereas disruptions of YCR047c and YDR140w lead to slow growth phenotypes.  
Normal growth was observed when the YDL201w, YDR465c, YHR209w, and YOR240w  
**genes** were disrupted. Initial analysis of protein methylation  
patterns of all mutants by amino acid analysis revealed that the YDR465c  
mutant has a defect in the methylation of the delta-nitrogen atom of  
arginine residues. We propose that YDR465c codes for the methyltransferase  
responsible for this recently characterized type of protein methylation,  
and we designate the enzyme as Rmt2 (protein **arginine  
methyltransferase**). In addition, we show that the methylation of  
susceptible residues in Rmt2 substrates is likely to take place on nascent  
polypeptide chains and that these substrates exist in the cell as fully  
methylated species. Interestingly, Rmt2 has 27% **sequence**  
identity over 138 amino acids to the mammalian guanidinoacetate  
N-methyltransferase, an enzyme responsible for methylating the  
delta-nitrogen of the small molecule guanidinoacetate.

L85 ANSWER 11 OF 76 MEDLINE DUPLICATE 5  
AB Hrp1p is a heterogeneous ribonucleoprotein (hnRNP) from the yeast  
*Saccharomyces cerevisiae* that is involved in the cleavage and  
polyadenylation of the 3'-end of mRNAs and mRNA export. In addition, Hrp1p  
is one of several RNA-binding proteins that are posttranslationally  
modified by methylation at arginine residues. By using functional  
recombinant Hrp1p, we have identified RNA **sequences** with  
specific high affinity binding sites. These sites correspond to the  
efficiency element for mRNA 3'-end formation, UAUUAUA. To examine the  
effect of methylation on specific RNA binding, purified recombinant  
**arginine methyltransferase** (Hmt1p) was used to methylate  
Hrp1p. Methylated Hrp1p binds with the same affinity to UAUUAUA-containing  
RNAs as unmethylated Hrp1p indicating that methylation does not affect  
specific RNA binding. However, RNA itself inhibits the methylation of  
Hrp1p and this inhibition is enhanced by RNAs that specifically bind Hrp1p.  
Taken together, these data support a model in which protein methylation  
occurs prior to protein-RNA binding in the nucleus.

L85 ANSWER 12 OF 76 SCISEARCH COPYRIGHT 2001 ISI (R)

L85 ANSWER 18 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V. DUPLICATE 9

AB Recombinant TIS21 protein was overexpressed in Escherichia coli harboring the expression vector plasmid pQE-30 carrying the TIS21 cDNA coding **sequence** containing an extra 120 nucleotides upstream. Employing this protein consisting of 158 amino acid residues of the main chain plus 40 residues of the fusion peptide, it was found that one of the protein methylase I group [S-adenosylmethionine:nuclear protein/histone-arginine N-methyltransferase; EC 2.1.1.23] methylated this protein. The methylation products were identified as guanidino-N-methylated arginines. Some of the kinetics of the reaction are described.

L85 ANSWER 24 OF 76 HCAPLUS COPYRIGHT 2001 ACS

AB This review, with 250 refs., summarizes the current status of protein arginine N-methylation reactions. These covalent modifications of proteins are now recognized in a no. of eukaryotic proteins and their functional significance is beginning to be understood. **Genes** that encode those methyltransferases specific for catalyzing the formation of asym. dimethylarginine have been identified. The enzyme modifies a no. of generally nuclear or nucleolar proteins that interact with nucleic acids, particularly RNA. Postulated roles for these reactions include signal transduction, nuclear transport, or a direct modulation of nucleic acid interactions. A second methyltransferase activity that sym. dimethylates an arginine residue in myelin basic protein, a major component of the axon sheath, has also been characterized. However, a **gene** encoding this activity has not been identified to date and the cellular function for this methylation reaction has not been clearly established. From the anal. of the **sequences** surrounding known arginine methylation sites, we have detd. consensus methyl-accepting **sequences** that may be useful in identifying novel substrates for these enzymes and may shed further light on their physiol. role. (c) 1998 Academic Press.

L85 ANSWER 33 OF 76 MEDLINE DUPLICATE 18

AB The intracytoplasmic domain (IC) of cytokine receptors provides docking sites for proteins which mediate signal transduction. Thus, in interferon-alpha,beta receptors (IFNAR1 and 2), the IC region binds protein-tyrosine and -serine/threonine kinases which phosphorylate the receptor and the associated Stat transcription factors. A two-hybrid screening was carried out to identify additional proteins which could interact with the IC domain of the IFNAR1 chain of the IFN-alpha,beta receptor. Several positive clones representing a protein **sequence** designated IR1B4 were recovered from a human cDNA library. IR1B4 was identified as the human homolog of PRMT1, a protein-**arginine methyltransferase** from rat cells. Flag-IR1B4 fusion proteins bind to the isolated IFNAR1 intracytoplasmic domain produced in Escherichia coli, as well as to the intact IFNAR1 chain extracted by detergent from human U266 cell membranes. S-Adenosylmethionine-dependent methyltransferase activity was precipitated by anti-IFNAR1 antibodies from untreated human cells. IR1B4/PRMT1 is involved in IFN action since U266 cells rendered deficient in this methyltransferase by antisense oligonucleotides become more resistant to growth inhibition by IFN. Methylation of proteins by enzymes which can attach to the IC domains of receptors may be a signaling mechanism complementing protein phosphorylation. Among substrates methylated by PRMT1 are RNA-binding heterogeneous nuclear ribonucleoproteins (hnRNPs) which are involved in mRNA processing, splicing and transport into the cytoplasm.

L85 ANSWER 34 OF 76 MEDLINE

DUPLICATE 19

AB The TIS21 immediate-early **gene** and leukemia-associated BTG1 **gene** encode proteins with similar **sequences**. Two-hybrid analysis identified a protein that interacts with TIS21 and BTG1. **Sequence** motifs associated with S-adenosyl-L-methionine binding suggested this protein might have methyltransferase activity. A glutathione S-transferase (GST) fusion of the putative methyltransferase modifies arginine residues, in appropriate protein substrates, to form NG-monomethyl and NG,NG-dimethylarginine (asymmetric). We term the protein- arginine N-methyltransferase (EC 2.1.1.23) **gene** "PRMT1, " for protein-**arginine methyltransferase 1**. GST-TIS21 and GST-BTG1 fusion proteins qualitatively and quantitatively modulate endogenous PRMT1 activity, using control and hypomethylated RAT1 cell extracts as methyl-accepting substrates. PRMT1 message appears ubiquitous, and is constitutive in mitogen-stimulated cells. Modulation of PRMT1 activity by transiently expressed regulatory subunits may be an additional mode of signal transduction following ligand stimulation.

L85 ANSWER 45 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.DUPLICATE 26

AB The enzymatic methylation of polypeptides on the guanidino group of internal arginine residues by S-adenosylmethionine:protein arginine N-methyltransferase (protein methylase I) yields N(G)-monomethylarginine, N(G),N(G)-dimethylarginine and N(G),N'(G)-dimethylarginine. It has commonly been observed that these arginine residues are present in glycine-and-arginine rich motifs. To understand structural features which are essential for serving as the methyl acceptor for protein methylase I, we have investigated substrate capacities of several synthetic oligopeptides whose **sequences** are homologous and/or analogous to the methyl acceptor region of the naturally occurring arginine-methylated proteins. These studies have led to the following conclusions. (i) The preferred amino-acid **sequence** of methyl-accepting peptides was shown to be an arginine-containing peptide with glycine in both the N- and C-flanking positions. While a tetrapeptide with such a **sequence** (residues 106-109 of bovine myelin basic protein) exhibited almost negligible substrate activity, an overlapping hexapeptide was a moderate substrate. (ii) Substitution of the C-flanking glycine in GKGRGL (residues 104-109 of myelin basic protein) with histidine, phenylalanine, lysine or aspartic acid completely abolished the ability of these hexapeptides to serve as substrates. (iii) A heptapeptide with a repeated glycine-arginine motif (GRGRGRG) was an excellent substrate for the enzyme. (iv) A cyclic octapeptide CGKGRGLC, which was formed by cyclization of GKGRGL by introduction of disulfide bridge to cross-link N- and C-terminus of the hexapeptide, was an even better substrate than the hexapeptide. (v) Upon HPLC amino-acid analysis, all enzymatically methyl-<sup>14</sup>C-labeled oligopeptides were found to yield predominantly N(G)-monomethylarginine with a minor fraction of N(G),N(G)-dimethylarginine in certain peptide samples. However, no N(G),N'(G)-dimethylarginine formation was detectable. (vi) The recombinant hnRNP protein A1 (residues 1-320) is known to be methylated at arginine-194 by nuclear-protein/histone protein methylase I (Rajpurohit et al. (1994) J. Biol. Chem. 269, 1079-1082). However, the hexapeptide (SSSQRG) which corresponds to residues 189-194 of protein A1 containing the methylatable arginine residue was relatively inert as a substrate. Furthermore, the N-terminal fragment of protein A1 (residues 1-196) generated by controlled trypsin digestion was also completely inactive as a substrate for the enzyme. These results indicate that the remainder of the A1 protein molecule plays an important though not yet understood role in enzymatic methylation of the arginine-194.

L85 ANSWER 47 OF 76 MEDLINE

DUPLICATE 27

AB Several proteins that interact with RNA, e.g. the heterogenous ribonucleoprotein particle A and B proteins, fibrillarin and nucleolin, contain the modified amino acid NG,NG-dimethylarginine. Here, we report that two synthetic peptides, Ac-GGRGGFGRGGFGRGGFG-NH<sub>2</sub> (R3) and GGFGRGGFG-NH<sub>2</sub> (R1), which are based on methylated **sequences** in fibrillarin and nucleolin, inhibit the methylation of a large majority of the methyl-accepting proteins observed in extracts of adenosine dialdehyde-treated PC12 cells. Concomitantly, the peptides themselves become methylated, suggesting that they compete for the same enzyme that carries out the bulk of N-methylation in PC12 cells. R3 potentially inhibits formation of NG,NG-dimethylarginine in PC12 substrates, with a lesser effect on NG-monomethylarginine and NG,N'-G-dimethylarginine. Bovine brain contains an activity that methylates PC12 methyl acceptors. After partial purification, the bovine methyltransferase efficiently modifies R3 and R1, yielding half maximal rates of methylation at approximately 0.2 and approximately 2 microM peptide, respectively. A search of the GenPept database for the FGGRGGF motif revealed 13 candidate methyl acceptors containing arginine and at most two similar substitutions or one mismatch. Of these, 10 are known or presumed to interact with RNA. These findings are consistent with the hypothesis that a majority of proteins containing NG,NG-dimethylarginine interact with RNA.

L85 ANSWER 76 OF 76 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.

AB S Adenosyl L ethionine was found to inhibit histone methylation by an **arginine methyltransferase** purified from rat liver cytoplasm, in the presence of S adenosyl L methionine. The enzyme is able to transfer alkyl groups from either of the activated amino acids to histone and with comparable efficiency. Histone methylation is also inhibited by N hydroxy 2 acetylaminofluorene (40% at 1 mM inhibitor concentration) and N acetoxy 2 acetylaminofluorene (69% at 1.0 mM inhibitor concentration). The **arginine methyltransferase** has a Km value of 4.4 .mu.M and is subject to product inhibition by S adenosyl L homocysteine (Ki = 1.2 .mu.M). Other similarities to the transfer ribonucleic acid methyltransferases are susceptibility to inhibition by adenine and the adenosine analogs, tubercidin and N 6 (.DELTA.2 isopentenyl) adenosine (16, 79, and 72%, respectively, at 1.0 mM inhibitor concentrations). The dyes ethidium bromide and acridine orange also inhibit the methyltransferase system (58 and 80%, respectively, at 1.0 mM inhibitor concentrations). The observation that arginine methylation in histones is inhibited by reactive metabolites of carcinogenic agents is used as a basis for proposing that interference in the normal histone methylation process, with resultant effects on fidelity of **gene** expression represents a preliminary step in carcinogenesis by such agents.

=> log y

COST IN U.S. DOLLARS

| SINCE FILE | TOTAL   |
|------------|---------|
| ENTRY      | SESSION |
| 40.18      | 165.11  |

FULL ESTIMATED COST

DISCOUNT AMOUNTS (FOR QUALIFYING ACCOUNTS)

| SINCE FILE | TOTAL   |
|------------|---------|
| ENTRY      | SESSION |
| -0.59      | -0.59   |

CA SUBSCRIBER PRICE

STN INTERNATIONAL LOGOFF AT 10:40:40 ON 21 JUN 2001